Diskretne strukture

Gašper Fijavž

2017

Univerza v Ljubljani
Fakulteta za računalništvo in informatiko

CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

51(075.8)

FIJAVŽ, Gašper
Diskretne strukture / Gašper Fijavž. - 2. izd. - Ljubljana : Založba FRI, 2017

ISBN 978-961-6209-90-8

289714432

1. elektronska izdaja knjige je na voljo na URL:
http://matematika.fri.uni-lj.si/ds/ds.pdf

Recenzenta: prof. dr. Bojan Orel, prof. dr. Marko Petkovšek
Založnik: Založba UL FRI, Ljubljana
Izdajatelj: UL Fakulteta za računalništvo in informatiko, Ljubljana
Oblikovanje platnic: prof. dr. Narvika Bovcon
Urednik: prof. dr. Franc Solina

Kazalo

Predgovor

Pred vami je učbenik predmeta *Diskretne strukture*, ki ga izvajamo na prvostopenjskem študiju Računalništva in informatike na Fakulteti za računalništvo in informatiko Univerze v Ljubljani.

Diskretne strukture so predmet tako univerzitetnega kot visokošolskega programa študija. Vsebine, obravnavane v tem učbeniku, pokrivajo oba predmeta.

V učbeniku uvedemo osnovne matematične strukture, ki jih računalničarji potrebujejo pri svojem nadaljnjem študiju.

Začnemo z osnovami matematične logike, nadaljujemo pa s teorijo množic, klasičnima začetnima stopnicama v matematiko. Z elementi teorije množic zgradimo nekaj naslednjih poglavij o relacijah, preslikavah in moči množic. Naslednji poglavji o teoriji števil in permutacijah lahko smatramo tudi kot uporabni verziji abstraktne algebre, čeprav se govoru o kolobarjih in grupah kot abstraktnih algebrskih strukturah skoraj popolnoma izognemo. Na koncu spoznamo dandanašnji vseprisotne grafe.

Aksiomatskemu pristopu k matematični logiki in teoriji množic se izognemo, delo je namenjeno študentom računalništva, uporabnikom (in ne graditeljem) matematičnih teorij. Vseeno pa brez matematične strogosti pri diskretnih strukturah ne gre. Veliko večino rezultatov želimo utemeljiti, izreki in trditve so praviloma opremljeni z ustreznimi dokazi.

Na tem mestu naj se najprej zahvalim Aleksandri Franc, Damirju Franetiču in Andreju Vodopivcu za pripombe v začetni fazi nastajanja te knjige. Za skrben pregled se zahvaljujem recenzentoma Bojanu Orlu in Marku Petkovšku. Uspela sta uloviti precej napak, ki sem jih nespretno posejal sam. Če je kakšna ostala, o tem (žal) ne dvomim, gre izključno na moj rovaš.

Učbenik je v celoti dostopen na spletnem naslovu

`matematika.fri.uni-lj.si/ds/ds.pdf`

Ljubljana, 2015 Gašper Fijavž

Poglavje 1

Izjavni račun

1.1 Izjave

Izjava je stavek, ki je *resničen* ali *neresničen* (in ne oboje hkrati), z eno besedo, ki ima *logično vrednost*.

Izjava

> *Ena in ena je dva.*

je resnična, pravimo tudi, da ima logično vrednost 1. Po drugi strani je izjava

> *Dva in pet je manjše kot tri.*

napačna, ima logično vrednost 0.

Že pri hitrem pregledu besedila opazimo, da vsi stavki niso izjave. Pri stavkih

> *Zapri vrata!*

in

> *Ko bi le jutri sijalo sonce.*

o logični vrednosti niti ne moremo razpravljati. S stališča slovnice se smemo omejiti na stavke v povednem naklonu. Toda tudi v tem primeru se lahko znajdemo v težavah. Stavek

> *Ta stavek ni resničen.*

namreč ni izjava. V nasprotnem primeru bi mu lahko določili logično vrednost. Toda premislimo lahko, da nobena od logičnih vrednosti zanj ni ustrezna. Če bi namreč bil resničen, potem bi moralo biti tisto, kar sam zase trdi, res. Da je neresničen. Če pa bi

bil neresničen, potem bi moral biti tudi resničen, saj domnevno o svoji logični vrednosti laže.

Stavka

> *Sonce sije.*

in

> *Tone skače po zeleni travi.*

sta izjavi, saj jima lahko določimo logični vrednosti. Ne gre pa brez težav. Logična vrednost prvega stavka je morda odvisna od kraja in časa. — Ali Sonce sije tudi ponoči ali če je oblačno? Je morda v času noči pri nas vsaj na južnem Pacifiku sončno vreme? Smo res prepričani, da je Sonce v času pred 5 milijardami let sijalo? — Temu problemu se lahko ognemo tako, da si predstavljamo situacijo (konkreten kraj, čas in tudi vreme), v katerem ima naš stavek natanko določeno logično vrednost.

V primeru drugega stavka je težava v imenu Tone. Na svetu je več ljudi z imenom Tone in vsi hkrati ne počnejo iste stvari. V težave ne zabredemo, če se odločimo, da ime, v tem primeru Tone, označuje natanko določenega fanta iz soseščine.

Izjave po zgradbi delimo na *enostavne* in *sestavljene*. Zgleda enostavnih izjav sta

> *Veter piha.*
>
> *Jure se vozi s kolesom.*

sestavljene izjave pa dobimo iz enostavn(ejš)ih z uporabo veznikov. Za primer navajam

> *Veter piha in Jure se vozi s kolesom.*
>
> *Ni res, da se Jure vozi s kolesom.*
>
> *Ker piha veter, se Jure ne vozi s kolesom.*

1.2 Izjavni vezniki

Izjavni vezniki predstavljajo način, kako enostavnejše izjave lepimo v bolj zapletene izjave. Pri tem želimo, da bo logična vrednost tako sestavljene izjave odvisna samo od izbire veznika in logičnih vrednosti sestavnih delov.

Z drugimi besedami, če eno resnično izjavo v sestavljeni izjavi nadomestimo z drugo resnično izjavo, se logična vrednost sestavljene izjave ne spremeni. Ravno tako se logična vrednost sestavljene izjave ne spremeni, če nadomestimo eno neresnično izjavo z drugo. To pomeni, da bomo izjavne veznike (tudi *logične veznike*, *izjavne povezave*) smeli definirati z uporabo pravilnostnih tabel.

V nadaljevanju razdelka bomo črki A in B uporabljali kot semantični spremenljivki, označevali bosta poljubni izjavi.

Negacijo izjave A označimo z $\neg A$ in beremo *"Ne A"*. Negacija je definirana z naslednjo pravilnostno tabelo.

A	$\neg A$
0	1
1	0

V levem stolpcu pravilnostne tabele imamo po posameznih vrsticah zapisane *vse* možne logične vrednosti izjave A, natanko dve sta. V desnem stolpcu sta zapisani ustrezni logični vrednosti negacije izjave A.

Negacija izjave A, $\neg A$, je resnična natanko tedaj, ko je izjava A neresnična.

Konjunkcijo izjav A in B označimo z $A \wedge B$ in beremo *"A in B"*. Konjunkcija izjav A in B je resnična natanko tedaj, ko sta obe izjavi A in B resnični, pravilnostna tabela konjunkcije je videti takole.

A	B	$A \wedge B$
0	0	0
0	1	0
1	0	0
1	1	1

Disjunkcijo izjav A in B označimo z $A \vee B$ in beremo *"A ali B"*. Disjunkcija izjav A in B je resnična natanko tedaj, ko je vsaj ena od izjav A oziroma B resnična. Oglejmo si njeno pravilnostno tabelo.

A	B	$A \vee B$
0	0	0
0	1	1
1	0	1
1	1	1

Implikacijo označimo z $A \Rightarrow B$ in beremo *"Iz A sledi B"* ali *"Če A, potem B"* ali tudi bolj formalno *"A implicira B"*. Implikacija $A \Rightarrow B$ je neresnična samo v primeru, ko je A resnična in B lažna izjava.

A	B	$A \Rightarrow B$
0	0	1
0	1	1
1	0	0
1	1	1

V implikaciji $A \Rightarrow B$ imata prvi člen A in drugi člen B različni vlogi. Prvemu členu implikacije pravimo tudi *antecédens*, drugemu členu implikacije pa *konsékvens*.

Ekvivalenco izjav A in B označimo z $A \Leftrightarrow B$ in jo preberemo *"A natanko tedaj, ko B"* ali *"A, če in samo če B"* ali celo *"A ekvivalentno B"*. Ekvivalenca $A \Leftrightarrow B$ je resnična natanko tedaj, ko imata A in B isto logično vrednost.

A	B	$A \Leftrightarrow B$
0	0	1
0	1	0
1	0	0
1	1	1

Ekskluzivno disjunkcijo izjav A in B označimo z $A \veebar B$ in beremo *"A ekskluzivni ali B"*. Ekskluzivna disjunkcija izjav A in B je resnična natanko tedaj, ko je *natančno* ena od izjav A ali B resnična, ali enakovredno, ko imata A in B različni logični vrednosti.

Pravilnostni tabeli disjunkcije in ekskluzivne disjunkcije se razlikujeta samo v primeru dveh resničnih členov.

A	B	$A \veebar B$
0	0	0
0	1	1
1	0	1
1	1	0

Izjavo, sestavljeno iz izjav A in B s *Shefferjevim veznikom*, označimo z $A \uparrow B$. Le-ta je neresnična samo v primeru, ko sta obe izjavi A in B resnični. Pravilnostna tabela $A \uparrow B$ ima ravno nasprotne vrednosti kot tabela konjunkcije $A \wedge B$.

A	B	$A \uparrow B$
0	0	1
0	1	1
1	0	1
1	1	0

Shefferjevemu izjavnemu vezniku pravimo tudi veznik NAND.

Izraz $A \downarrow B$ označuje izjavo, sestavljeno iz A in B z uporabo *Peirce-Łukasiewiczevega veznika*, tudi veznika NOR. Izjava $A \downarrow B$ je resnična natanko tedaj, ko sta obe izjavi A in B lažni. Veznik *niti-niti* je naravna ustreznica logičnega veznika $\downarrow$, ki ga opišemo z naslednjo pravilnostno tabelo.

A	B	$A \downarrow B$
0	0	1
0	1	0
1	0	0
1	1	0

1.2.1 Še več izjavnih veznikov

Definirali smo sedem dvomestnih izjavnih veznikov, ki jih lahko samo deloma motiviramo v naravnem jeziku. Vsekakor bomo uspeli v proznem tekstu prepoznati stavke, ki so po zgradbi konjunkcije, implikacije, včasih celo disjunkcije. Ekvivalence pogosto zasledimo v tehničnih tekstih, denimo v matematični literaturi, ravno tako ekskluzivne disjunkcije. Shefferjev in Peirce-Łukasiewiczev veznik sta pomembna s teoretičnega vidika, kar bomo spoznali v enem od naslednjih razdelkov.

Bi lahko definirali še več izjavnih veznikov? Vsekakor. Izjavne veznike definiramo s pomočjo pravilnostnih tabel in obstaja kar 16 različnih dvomestnih izjavnih veznikov. Seveda moremo definirati tudi večmestne izjavne veznike, ki jih je posledično ustrezno več — število n-mestnih izjavnih veznikov je enako 2^{2^n}.

1.3 Izjavni izrazi

Izjavne izraze definiramo induktivno — (I1) in (I2) opisujeta bazične izjavne izraze, tiste najenostavnejše. Konstrukcijsko pravilo (I3) pa opisuje, kako zgradimo nove izjavne izraze iz že konstruiranih:

(I1) *Izjavni konstanti* 0 in 1, ki jima pravimo tudi *laž* in *resnica*, sta izjavna izraza.

(I2) *Izjavne spremenljivke* p, q, r, ... so izjavni izrazi.

(I3) Če je F n-mestni izjavni veznik in so A_1, A_2, ..., A_n izjavni izrazi, potem je tudi

$$F(A_1, \ldots, A_n)$$

izjavni izraz.

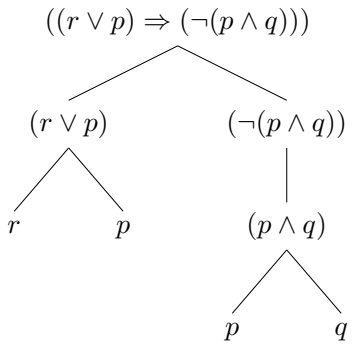
Pravilo (I3) je treba interpretirati kot vzorec, z uporabo različnih izjavnih veznikov lahko (I3) preberemo na veliko različnih načinov. V kar največji meri bomo uporabljali le enomestne in dvomestne izjavne veznike. Če se omejimo le na tiste, ki smo jih spoznali v prejšnjem razdelku, smemo pravilo (I3) nadomestiti s pravili (I3a) in (I3b):

(I3a) Če je A izjavni izraz, potem je tudi $(\neg A)$ izjavni izraz.

(I3b) Če sta A in B izjavna izraza, potem so tudi

$$(A \wedge B),\ (A \vee B),\ (A \Rightarrow B),\ (A \Leftrightarrow B),\ (A \veebar B),\ (A \uparrow B) \text{ in } (A \downarrow B)$$

izjavni izrazi.



Slika 1.1: Konstrukcijsko drevo izraza $((r \vee p) \Rightarrow (\neg(p \wedge q)))$

1.3.1 Konstrukcijsko drevo

Induktivni opis izjavnih izrazov porodi *konstrukcijsko drevo*, ki ga definiramo rekurzivno. Konstrukcijska drevesa izjavnih spremenljivk in izjavnih konstant vsebujejo eno samo vozlišče, ki je enako takšni spremenljivki oziroma konstanti. Če za izjavni izraz A velja $A = F(A_1, \ldots, A_n)$, potem izraz A postavimo v koren njegovega konstrukcijskega drevesa, sinovi korena pa so konstrukcijska drevesa izjavnih izrazov $A_1, \ldots, A_n$.

Morda nas zmoti pretirana uporaba oklepajev, vsaj zunanji par oklepajev se res zdi nepotreben. Z dogovorom o *prednosti izjavnih veznikov* se lahko izognemo še dodatnim oklepajskim parom.

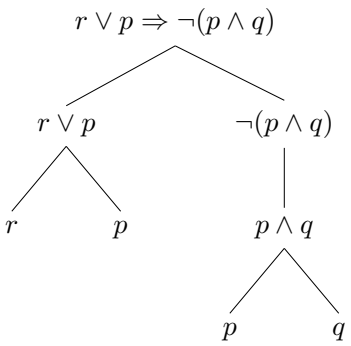
Izjavne veznike po prednosti razdelimo v pet razredov: negacija $\neg$ veže najmočneje, takoj za njo so po prednosti enakovredni konjunkcija $\wedge$, Shefferjev veznik $\uparrow$ in Peirce-Łukasiewiczev veznik $\downarrow$. Naslednji par po prednosti enakovrednih veznikov sta disjunkcija $\vee$ in ekskluzivna disjunkcija $\veebar$, ki vežeta močneje kot implikacija $\Rightarrow$, le-ta pa spet močneje kot ekvivalenca $\Leftrightarrow$.

$\neg$	$\wedge, \uparrow, \downarrow$	$\vee, \veebar$	$\Rightarrow$	$\Leftrightarrow$

Tabela 1.1: Izjavni vezniki razvrščeni po prednosti.

V primeru, ko v izrazu nastopa več po prednosti enakovrednih izjavnih veznikov, jih združujemo z leve proti desni.

Izjavna izraza I in I', ki se razlikujeta samo v nekaterih parih (v skladu z zgornjim dogovorom) nepotrebnih oklepajev, smemo *enačiti*, pravimo, da sta izraza I in I' *enaka*.



Slika 1.2: Odpravljeni oklepaji v konstrukcijskem drevesu izraza $r \vee p \Rightarrow \neg(p \wedge q)$

Oglejmo si nekaj zgledov parov enakih izjavnih izrazov.

$$\begin{aligned} p \vee q \Rightarrow r \Leftrightarrow s \wedge t &= ((p \vee q) \Rightarrow r) \Leftrightarrow (s \wedge t) \\ p \vee q \veebar r \vee p &= ((p \vee q) \veebar r) \vee p \\ p \uparrow q \wedge r \downarrow q &= ((p \uparrow q) \wedge r) \downarrow q \\ p \wedge q \vee p \wedge r &= (p \wedge q) \vee (p \wedge r) \end{aligned}$$

Na levi strani so predstavljeni izjavni izrazi brez zapisanih oklepajev, izrazi na desni pa so zapisani z, v skladu z dogovorom, odvečnimi oklepaji. V nekaterih primerih bomo zaradi večje preglednosti katerega od načeloma odvečnih parov oklepajev ohranili.

Naslednja izjavna izraza nista enaka.

$$p \vee q \wedge r \quad \neq \quad (p \vee q) \wedge r.$$

Konjunkcija veže močneje kot disjunkcija, izraz $p \vee q \wedge r$ smemo enačiti z izrazom $p \vee (q \wedge r)$.

S pomočjo konstrukcijskega drevesa moremo definirati tudi pojem nastopanja enega izjavnega izraza v drugem. Pravimo, da izjavni izraz I *nastopa* v izjavnem izrazu J, če je izraz I vozlišče konstrukcijskega drevesa izraza J. Tako v izrazu $I = r \vee p \Rightarrow \neg(p \wedge q)$ nastopajo naslednji izjavni izrazi

$$p, \quad q, \quad r, \quad r \vee p, \quad p \wedge q, \quad \neg(p \wedge q) \quad \text{in tudi} \quad r \vee p \Rightarrow \neg(p \wedge q),$$

medtem ko izjavni izraz $p \Rightarrow \neg(p \wedge q)$ v izrazu I ne nastopa.

Za konec definirajmo še mere za velikost izjavnega izraza. *Globina* izjavnega izraza I je enaka dolžini najdaljše poti do korenskega vozlišča v konstrukcijskem drevesu izraza I. *Dolžina* izjavnega izraza je vsota števila listov konstrukcijskega drevesa in števila uporabljenih logičnih veznikov. Vsak uporabljeni izjavni veznik ustreza natanko enemu notranjemu vozlišču konstrukcijskega drevesa, torej je dolžina izjavnega izraza enaka številu vseh vozlišč konstrukcijskega drevesa.

Izjavni izraz $r \vee p \Rightarrow \neg(p \wedge q)$, predstavljen na sliki 1.2, ima globino enako 3 in dolžino enako 8.

1.3.2 Pravilnostna tabela

Pravilnostna tabela (tudi *resničnostna tabela*) izjavnega izraza I je podatkovna struktura (zapišemo jo v tabelarični obliki), ki za vsak nabor logičnih vrednosti spremenljivk, ki nastopajo v I, podaja logično vrednost, ki jo dobimo, če spremenljivke izraza I nadomestimo z logičnimi vrednostmi ustreznega nabora.

Pravilnostno tabelo izjavnega izraza I izračunamo rekurzivno — če poznamo pravilnostne tabele vseh izjavnih izrazov, ki nastopajo v I (z izjemo samega izraza I), potem pravilnostno tabelo izraza I določimo z uporabo ustrezne pravilnostne tabele zadnjega uporabljenega logičnega veznika.

Pravilnostno tabelo izjavnega izraza $I = r \vee p \Rightarrow \neg(p \wedge q)$ smo zapisali v tabelo 1.2. Pravilnostne tabele izrazov

$$p \Rightarrow (q \Rightarrow p), \quad (p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p) \quad \text{in} \quad (p \Rightarrow (q \Rightarrow r)) \wedge (p \Rightarrow q) \wedge p \wedge \neg r,$$

pa najdemo v tabelah 1.3 in 1.4.

p	q	r	$r \vee p$	$p \wedge q$	$\neg(p \wedge q)$	$r \vee p \Rightarrow \neg(p \wedge q)$
0	0	0	0	0	1	1
0	0	1	1	0	1	1
0	1	0	0	0	1	1
0	1	1	1	0	1	1
1	0	0	1	0	1	1
1	0	1	1	0	1	1
1	1	0	1	1	0	0
1	1	1	1	1	0	0

Tabela 1.2: Pravilnostna tabela izraza $r \vee p \Rightarrow \neg(p \wedge q)$, skupaj s pravilnostnimi tabelami izrazov, ki v njem nastopajo.

Opazimo, da imata izraza $p \Rightarrow (q \Rightarrow p)$ in $(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$ pri vseh možnih naborih logičnih vrednosti svojih spremenljivk (vedno, bi rekli po domače) logično vrednost enako 1. Takšnemu izjavnemu izrazu pravimo *tavtologija*. Tipični zgledi tavtologij so tudi logična konstanta 1 ter izrazi $p \vee \neg p$, $p \Rightarrow p$, $p \Leftrightarrow p$.

Protislovje je izjavni izraz, ki ima pri vseh možnih naborih logičnih vrednosti spremenljivk logično vrednost enako 0. Izjavni izraz $(p \Rightarrow (q \Rightarrow r)) \wedge (p \Rightarrow q) \wedge p \wedge \neg r$ je protislovje, nadaljnji zgledi so tudi 0, $p \wedge \neg p$, $\neg p \Leftrightarrow p$. Izrazu, ki ni niti tavtologija niti protislovje, pravimo *nevtralni izjavni izraz*. V pravilnostni tabeli nevtralnega izjavnega izraza najdemo tako ničle kot enice. Velika večina izjavnih izrazov je nevtralnih.

p	q	$p \Rightarrow (q \Rightarrow p)$	$(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$
0	0	1	1
0	1	1	1
1	0	1	1
1	1	1	1

Tabela 1.3: Tavtologiji $p \Rightarrow (q \Rightarrow p)$ in $(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$.

p	q	r	$(p \Rightarrow (q \Rightarrow r)) \wedge (p \Rightarrow q) \wedge p \wedge \neg r$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	0

Tabela 1.4: Protislovje $(p \Rightarrow (q \Rightarrow r)) \wedge (p \Rightarrow q) \wedge p \wedge \neg r$.

1.4 Enakovredni izjavni izrazi

Izračunajmo pravilnostni tabeli izjavnih izrazov $(\neg p \vee q) \wedge r$ in $\neg(r \Rightarrow \neg(p \Rightarrow q))$, glej tabelo 1.5. Opazimo, da imata izjavna izraza v vsaki vrstici pravilnostne tabele isto logično vrednost.

Za izjavna izraza I in J pravimo, da sta *enakovredna*, pišemo tudi $I \sim J$, če imata za vsak nabor logičnih vrednosti spremenljivk isto logično vrednost.

Če za izjavna izraza I in J obstaja nabor logičnih vrednosti spremenljivk, pri katerih imata različni logični vrednosti, potem izraza I in J *nista* enakovredna. Če je za uteme-

p	q	r	$(\neg p \vee q) \wedge r$	$\neg(r \Rightarrow \neg(p \Rightarrow q))$
0	0	0	0	0
0	0	1	1	1
0	1	0	0	0
0	1	1	1	1
1	0	0	0	0
1	0	1	0	0
1	1	0	0	0
1	1	1	1	1

Tabela 1.5: Enakovredna izjavna izraza $(\neg p \vee q) \wedge r$ in $\neg(r \Rightarrow \neg(p \Rightarrow q))$.

ljitev enakovrednosti izrazov potrebno preveriti celotni njuni pravilnostni tabeli (število vrstic pravilnostne tabele narašča eksponentno v odvisnosti od števila spremenljivk), je za dokazovanje nasprotne trditve, da izraza nista enakovredna, dovolj poiskati en sam nabor vrednosti (lahko bi rekli tudi eno samo vrstico pravilnostne tabele), v kateri se njuni logični vrednosti ne ujemata.

Tako, denimo, izraza $(p \Rightarrow q) \Rightarrow r$ in $p \Rightarrow (q \Rightarrow r)$ nista enakovredna, kar preverimo z vstavljanjem logičnih vrednosti 0,1,0 namesto spremenljivk p, q, r:

p	q	r	$(p \Rightarrow q) \Rightarrow r$	$p \Rightarrow (q \Rightarrow r)$
0	1	0	0	1

Enakovrednost izjavnih izrazov je v tesni zvezi z veznikom ekvivalence:

Izrek 1.1 *Izjavna izraza I in J sta enakovredna natanko tedaj, ko je izraz $I \Leftrightarrow J$ tavtologija.*

Dokaz. Izraz $I \Leftrightarrow J$ je tavtologija natanko tedaj, ko je resničen pri vseh možnih naborih logičnih vrednosti spremenljivk. Če upoštevamo opis logičnega veznika $\Leftrightarrow$, se to zgodi natanko tedaj, ko imata izraza I in J vedno, to je pri vseh možnih naborih logičnih vrednosti spremenljivk, isto logično vrednost. To pa je natančno opis enakovrednosti $I \sim J$. □

Z opazovanjem pravilnostnih tabel lahko za enakovrednost izjavnih izrazov pokažemo tudi naslednje lastnosti.

Izrek 1.2 *Za poljubne izjavne izraze I_1, I_2 in I_3 velja*

(i) $I_1 \sim I_1$,

(ii) *če je* $I_1 \sim I_2$, *potem je* $I_2 \sim I_1$, *in*

(iii) *če je* $I_1 \sim I_2$ *in* $I_2 \sim I_3$, *potem je* $I_1 \sim I_3$.

□

Z izrekom 1.2 smo dejansko pokazali, da je enakovrednost $\sim$ ekvivalenčna[1] relacija v družini vseh izjavnih izrazov. Obenem nam isti izrek omogoča tudi dokaz enakovrednosti izjavnih izrazov z *izpeljavo*.

Dokaz enakovrednosti $I \sim J$ lahko zapišemo kot zaporedje izjavnih izrazov

$$I_1,\ I_2,\ I_3,\ \ldots, I_{n-1},\ I_n \tag{1.1}$$

z začetnim členom I_1 enakim I in zadnjim členom I_n enakim J, v katerem za vsak par zaporednih členov I_i in I_{i+1} velja enakovrednost $I_i \sim I_{i+1}$. Z rutinsko uporabo indukcije se lahko prepričamo v enakovrednost $I \sim J$.

[1] Relacije, tudi ekvivalenčne, obdelamo v poglavju 4.

Kako pa v zaporedju (1.1) preverimo enakovrednosti $I_i \sim I_{i+1}$ za $i = 1, \ldots, n-1$? Na voljo bomo imeli dve orodji, substitucijo in zakone izjavnega računa.

Izrek 1.3 *Denimo, da izjavni izraz I_1 nastopa v izrazu J_1, pišemo lahko tudi $J_1 = F(I_1)$. Z zamenjavo vstopa izraza I_1 z izrazom I_2 pridelamo izraz $J_2 = F(I_2)$. Če sta izraza I_1 in I_2 enakovredna, potem sta enakovredna tudi izraza J_1 in J_2.*

Dokaz. Pravilnostni tabeli izrazov J_1 in J_2 sta odvisni samo od pravilnostnih tabel izjavnih izrazov, ki nastopajo v J_1 oziroma J_2. Ker sta I_1 in I_2 enakovredna, imata enaki pravilnostni tabeli. Posledično sta enaki tudi pravilnostni tabeli izrazov J_1 in J_2. □

Z uporabo izreka 1.3 in ostrega pogleda v smeri tabele 1.5 se lahko prepričamo, da sta izjavna izraza

$$(\neg p \vee q) \wedge r \Rightarrow r \wedge p \quad \text{in} \quad \neg(r \Rightarrow \neg(p \Rightarrow q)) \Rightarrow r \wedge p \tag{1.2}$$

enakovredna, saj sta enakovredna izraza

$$(\neg p \vee q) \wedge r \quad \text{in} \quad \neg(r \Rightarrow \neg(p \Rightarrow q)),$$

ki v prvotnih izrazih (1.2) nastopata na isti način.

1.5 Zakoni izjavnega računa

Zakoni izjavnega računa so nekateri pomembni pari enakovrednih tipov izjavnih izrazov. Do konca razdelka izraze A, B, C interpretiramo kot semantične spremenljivke[2], nadomestimo jih lahko s katerimikoli izjavnimi izrazi. Z uporabo izreka 1.3 bi tudi zakone izjavnega računa smeli navajati zapisane z izjavnimi spremenljivkami.

(1) Zakon dvojne negacije:

$$\neg\neg A \sim A$$

(2) Idempotenca:

$$A \wedge A \sim A \qquad A \vee A \sim A$$

(3) Komutativnost:

$$\begin{aligned} A \wedge B &\sim B \wedge A & A \vee B &\sim B \vee A \\ A \Leftrightarrow B &\sim B \Leftrightarrow A & A \veebar B &\sim B \veebar A \\ A \uparrow B &\sim B \uparrow A & A \downarrow B &\sim B \downarrow A \end{aligned}$$

[2]Vse vstope črke A nadomestimo z istim izjavnim izrazom, ki pa sme biti poljuben. Tudi za črki B in tudi C smemo postopati enako.

(4) Asociativnost:

$$(A \wedge B) \wedge C \sim A \wedge (B \wedge C) \qquad (A \vee B) \vee C \sim A \vee (B \vee C)$$
$$(A \Leftrightarrow B) \Leftrightarrow C \sim A \Leftrightarrow (B \Leftrightarrow C) \qquad (A \veebar B) \veebar C \sim A \veebar (B \veebar C)$$

(5) Absorpcija:

$$A \wedge (A \vee B) \sim A \qquad A \vee (A \wedge B) \sim A$$

(6) Distributivnost:

$$(A \vee B) \wedge C \sim (A \wedge C) \vee (B \wedge C)$$
$$(A \wedge B) \vee C \sim (A \vee C) \wedge (B \vee C)$$
$$(A \veebar B) \wedge C \sim (A \wedge C) \veebar (B \wedge C)$$

(7) de Morganova zakona:

$$\neg(A \vee B) \sim \neg A \wedge \neg B \qquad \neg(A \wedge B) \sim \neg A \vee \neg B$$

(8) Kontrapozicija:

$$A \Rightarrow B \sim \neg B \Rightarrow \neg A$$

(9) Tavtologija in protislovje, 0 in 1:

$$A \Rightarrow A \sim 1 \qquad A \Leftrightarrow A \sim 1$$
$$A \vee \neg A \sim 1 \qquad A \wedge \neg A \sim 0$$

(10) Substitucija 0 in 1:

$$A \wedge 0 \sim 0 \qquad A \vee 0 \sim A$$
$$A \wedge 1 \sim A \qquad A \vee 1 \sim 1$$
$$A \Rightarrow 0 \sim \neg A \qquad 0 \Rightarrow A \sim 1$$
$$A \Rightarrow 1 \sim 1 \qquad 1 \Rightarrow A \sim A$$

(11) Lastnosti implikacije:

$$A \Rightarrow B \sim \neg A \vee B \qquad \neg(A \Rightarrow B) \sim A \wedge \neg B$$

(12) Lastnosti ekvivalence:

$$A \Leftrightarrow B \sim (A \Rightarrow B) \wedge (B \Rightarrow A) \qquad A \Leftrightarrow B \sim (A \wedge B) \vee (\neg A \wedge \neg B)$$
$$\neg(A \Leftrightarrow B) \sim \neg A \Leftrightarrow B$$

(13) Lastnosti Shefferjevega in Pierce-Łukasiewiczevega veznika:

$$A \uparrow B \ \sim \ \neg(A \wedge B) \qquad A \downarrow B \ \sim \ \neg(A \vee B)$$

Vse zgoraj omenjene zakone izjavnega računa bi bilo potrebno dokazati, morda z uporabo pravilnostnih tabelic. Ker gre za relativno enostavne, a dolgotrajne račune, bomo zadevo izpustili.

Vseeno zakoni izjavnega računa ne bodo ostali brez komentarja. Hiter pregled zakonov prikaže, da zakoni v zvezi z negacijo in disjunkcijo nastopajo v parih, tako poznamo dva distributivnostna zakona, ki povezujeta veznika $\wedge$ in $\vee$, dve varianti de Morganovih zakonov in dve vrsti absorpcije. Opažanje lahko strnemo v naslednji izrek:

Izrek 1.4 (dualnost med $\wedge$ in $\vee$) *Naj bo $I = F(0, 1, \wedge, \vee, p_1, \ldots, p_k)$ izjavni izraz, v katerem poleg izjavnih spremenljivk $p_1, \ldots, p_k$ in negacije $\neg$ (lahko) nastopata tudi izjavni konstanti 0 in 1, uporabljamo pa lahko tudi izjavna veznika $\wedge$ in $\vee$.*

Denimo, da izjavni izraz $\tau(I) = F(1, 0, \vee, \wedge, \neg p_1, \ldots, \neg p_k)$ pridelamo iz izjavnega izraza I_1 tako, da

- *vse vstope logične konstante 0 (oziroma 1) nadomestimo s konstanto 1 (oziroma 0),*
- *zamenjamo vsako uporabo $\wedge$ (oziroma $\vee$) z $\vee$ (oziroma $\wedge$) in*
- *vsako od spremenljivk nadomestimo z njeno negacijo.*

Potem velja

$$\tau(I) \ \sim \ \neg I.$$

Dokaz. Dokazujemo z indukcijo po globini n izjavnega izraza I.

Baza indukcije ($n = 0$) obravnava primer, ko je I bodisi izjavna spremenljivka bodisi logična konstanta. Ločimo tri primere.

a) $I = p_i$ za neki $i \in \{1, \ldots, k\}$: $\tau(I) = \neg p_i = \neg I$

b) $I = 0$: $\tau(I) = 1 \ \sim \ \neg 0 = \neg I$

c) $I = 1$: $\tau(I) = 0 \ \sim \ \neg 1 = \neg I$

Indukcijski korak obravnava primer, ko je globina izjavnega izraza I vsaj 1 ($n \geq 1$). V tem primeru I vsebuje vsaj en izjavni veznik, ki veže enega ali več podizrazov globine $< n$, za katere po indukcijski predpostavki (i.p.) izrek že velja. Glede na zadnji izjavni veznik, ki ga uporabimo pri konstrukciji izraza I, znova ločimo tri primere.

a) $I = \neg I_1$:

$$\tau(I) = \tau(\neg I_1) \overset{\text{def.}\tau}{=} \neg\tau(I_1) \overset{\text{i.p.}}{\sim} \neg\neg I_1 = \neg I$$

b) $I = I_1 \vee I_2$:

$$\tau(I) = \tau(I_1 \vee I_2) \overset{\text{def.}\tau}{=} \tau(I_1) \wedge \tau(I_2) \overset{\text{i.p.}}{\sim} \neg I_1 \wedge \neg I_2 \ \sim \ \neg(I_1 \vee I_2) = \neg I$$

c) $I = I_1 \wedge I_2$:

$$\tau(I) = \tau(I_1 \wedge I_2) \overset{\text{def.}\tau}{=} \tau(I_1) \vee \tau(I_2) \overset{\text{i.p.}}{\sim} \neg I_1 \vee \neg I_2 \ \sim \ \neg(I_1 \wedge I_2) = \neg I$$

□

Asociativnost ekskluzivne disjunkcije še zdaleč ni očitna (za razliko od asociativnosti konjunkcije in disjunkcije, ki ju je relativno enostavno opravičiti). Naslednji izrek pa nam bo razkril še več, saj iz njega poleg asociativnosti sledi tudi komutativnost veznika $\underline{\vee}$.

Izrek 1.5 *Naj bo I izjavni izraz, ki ga dobimo z vezavo izjavnih izrazov $I_1, I_2, \ldots, I_n$ z uporabo ekskluzivne disjunkcije. Potem je I resničen natanko tedaj, ko je natanko liho mnogo členov izmed $I_1, \ldots, I_n$ resničnih.*

Preden se lotimo dokaza, izrek še komentirajmo. Izrazi $I_1, \ldots, I_n$ niso nujno različni, natančneje, če med členi I_i isti člen nastopa večkrat, ga moramo ustrezno mnogokrat navesti v zaporedju $I_1, \ldots, I_n$. Iz izreka enostavno sledijo, kot primer, naslednje enakovrednosti

$$(p \underline{\vee} q) \underline{\vee} (r \underline{\vee} s) \ \sim \ p \underline{\vee} ((q \underline{\vee} r) \underline{\vee} s) \qquad p \underline{\vee} p \underline{\vee} p \ \sim \ p$$
$$p \underline{\vee} p \underline{\vee} p \underline{\vee} p \ \sim \ 0 \qquad p \underline{\vee} q \underline{\vee} p \underline{\vee} q \underline{\vee} q \ \sim \ q$$

Dokaz. Dokazujemo z indukcijo po številu členov ekskluzivne disjunkcije. Trditev očitno velja za $n = 1$. V primeru, ko je $n = 2$, je $I = I_1 \underline{\vee} I_2$ in je I resničen natanko tedaj, ko imata I_1 in I_2 različni logični vrednosti. To pa se zgodi samo v primeru, ko je natančno eden (in s tem liho mnogo) izmed členov I_1, I_2 resničnih.

Indukcijsko predpostavko uporabimo v krepkem smislu. Privzamemo, da je za vsako ekskluzivno disjunkcijo s strogo manj kot n členi trditev izreka pravilna. Od tod tudi sledi, da položaj oklepajev ne vpliva na logično vrednost tako kratke ekskluzivne disjunkcije.

Pri ustrezno izbranem indeksu i, $2 \leq i \leq n$, je izraz I enakovreden

$$(I_1 \underline{\vee} \ldots \underline{\vee} I_{i-1}) \underline{\vee} (I_i \underline{\vee} \ldots \underline{\vee} I_n).$$

Po definiciji ekskluzivne disjunkcije je I resničen natanko tedaj, ko imata člena $I_1 \underline{\vee} \ldots \underline{\vee} I_{i-1}$ in $I_i \underline{\vee} \ldots \underline{\vee} I_n$ različni logični vrednosti. To se po indukcijski predpostavki zgodi natanko tedaj, ko sta števili pravilnih členov v ekskluzivnih disjunkcijah $I_1 \underline{\vee} \ldots \underline{\vee} I_{i-1}$ in $I_i \underline{\vee} \ldots \underline{\vee} I_n$ različne parnosti. Odtod pa sledi, da je natanko liho mnogo členov v zaporedju $I_1, \ldots, I_n$ resničnih. □

1.6 Normalni obliki in polni nabori izjavnih veznikov

V začetku razdelka navedimo nekaj nujno potrebnih definicij. *Literal* je krajše ime za izjavno spremenljivko oziroma njeno negacijo. Primeri literalov so p, $\neg p$, q, $\neg q$. *Osnovna konjunkcija* je bodisi osamljen literal ali pa konjunkcija večjega števila literalov. Zgledi osnovnih konjunkcij so $p \wedge q$, $\neg p \wedge q \wedge r$, pa tudi $p \wedge p$ in $\neg q$.

Osnovno disjunkcijo definiramo analogno: osnovna disjunkcija je bodisi literal bodisi disjunkcija večjega števila literalov, $p \vee q$, $\neg p \vee q \vee r$ in p so zgledi osnovnih disjunkcij.

Nadaljujmo z nalogo. Zapisati želimo izjavni izraz I, ki ima naslednjo pravilnostno tabelo:

p	q	r	I
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

Razmišljamo takole:

> *I je resničen*

natanko tedaj, ko

> *smo v 1. vrstici ali*
> *smo v 2. vrstici ali*
> *smo v 3. vrstici ali*
> *smo v 6. vrstici ali*
> *smo v 8. vrstici resničnostne tabele.*

To je spet enakovredno dejstvu, da

> *je p lažen in q lažen in r lažen ali*
> *je p lažen in q lažen in r resničen ali*
> *je p lažen in q resničen in r lažen ali*
> *je p resničen in q lažen in r resničen ali*
> *je p resničen in q resničen in r resničen,*

kar končno lahko prepišemo v izjavni izraz

$$(\neg p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r) \vee (\neg p \wedge q \wedge \neg r) \vee (p \wedge \neg q \wedge r) \vee (p \wedge q \wedge r)$$

Izjavni izraz, ki smo ga konstruirali po tem *receptu*, ima prav posebno strukturo: gre za disjunkcijo osnovnih konjunkcij. Še več, v osnovnih konjunkcijah nastopajo natančno vse tri spremenljivke iz resničnostne tabele, vsaka osnovna konjunkcija vsebuje natanko tri literale.

Naj bo I izjavni izraz. *Disjunktivna normalna oblika* izjavnega izraza I je izjavni izraz I_{DNO}, za katerega velja

(DNO1) $I_{DNO} \sim I$ in

(DNO2) I_{DNO} je disjunkcija osnovnih konjunkcij,

pri čemer dovolimo tudi disjunkcije z enim samim členom.

Disjunktivna oblika izjavnega izraza I je *polna*, če vsaka izmed osnovnih konjunkcij vsebuje vse izjavne spremenljivke, ki nastopajo v I, vsako bodisi kot spremenljivko bodisi kot njeno negacijo.

Disjunktivna normalna oblika izjavnega izraza ni enolično določena. Izjavni izraz $p \Rightarrow q$ lahko enakovredno prepišemo kot $\neg p \vee q$, zapišemo lahko pa tudi njeno polno disjunktivno normalno obliko $(\neg p \wedge \neg q) \vee (\neg p \wedge q) \vee (p \wedge q)$.

Trditev 1.6 *Vsak izjavni izraz I lahko enakovredno zapišemo v disjunktivni normalni obliki. Če I ni protislovje, potem lahko I enakovredno zapišeno v polni disjunktivni normalni obliki.*

Polno disjunktivno normalno obliko izjavnega izraza I, ki ni protislovje, konstruiramo tako, da za vsak nabor logičnih vrednosti spremenljivk, pri katerih je I resničen, pripravimo eno osnovno konjunkcijo, v kateri nastopajo v tem naboru resnične spremenljivke in negacije v tem naboru lažnih spremenljivk.

Dokaz. Protislovje lahko zapišemo kot $p \wedge \neg p$, disjunkcijo z enim samim členom. Če I ni protislovje, potem lahko z uporabo recepta zapišemo celo polno disjunktivno obliko izjavnega izraza I. □

Prvotno nalogo, konstrukcijo izjavnega izraza s predpisano pravilnostno tabelo, bi lahko zastavili tudi drugače.

I je resničen

natanko tedaj, ko

nismo v 4. vrstici in

nismo v 5. vrstici in

nismo v 7. vrstici resničnostne tabele.

To je enakovredno dejstvu, da

je p resničen ali q lažen ali r lažen in

je p lažen ali q resničen ali r resničen in

je p lažen ali q lažen ali r resničen,

kar prepišemo v izjavni izraz

$$(p \vee \neg q \vee \neg r) \ \wedge \ (\neg p \vee q \vee r) \ \wedge \ (\neg p \vee \neg q \vee r).$$

Tokratna verzija recepta je dualna prejšnjemu receptu. Posledično je zgradba dobljenega izraza dualna (v smislu zamenjave vlog konjunkcije in disjunkcije), kar nas motivira za naslednjo definicijo.

Konjunktivna normalna oblika izjavnega izraza I je izjavni izraz I_{KNO}, za katerega velja

(KNO1) $I_{KNO} \ \sim \ I$ in

(KNO2) I_{KNO} je konjunkcija osnovnih disjunkcij,

pri čemer znova dovolimo tudi konjunkcije z enim samim členom.

Konjunktivna normalna oblika izjavnega izraza I je *polna*, če vsaka izmed osnovnih disjunkcij vsebuje vse izjavne spremenljivke, ki nastopajo v I, vsako bodisi kot spremenljivko bodisi kot njeno negacijo.

Tudi konjunktivna normalna oblika izjavnega izraza ni enolično določena.

Trditev 1.7 *Vsak izjavni izraz I lahko enakovredno zapišemo v konjunktivni normalni obliki. Če I ni tavtologija, potem lahko I enakovredno zapišemo v polni konjunktivni normalni obliki.*

Polno konjunktivno normalno obliko izjavnega izraza I, ki ni tavtologija, konstruiramo tako, da za vsak nabor logičnih vrednosti spremenljivk, pri katerih je I neresničen, pripravimo eno osnovno disjunkcijo, v kateri nastopajo v tem naboru lažne spremenljivke in negacije v tem naboru resničnih spremenljivk.

Dokaz. Tavtologijo lahko zapišemo kot $p \vee \neg p$. Če izraz I ni tavtologija, ga z uporabo gornjega postopka enakovredno prepišemo v polni konjunktivni normalni obliki. □

Zgodba o disjunktivni in konjunktivni normalni obliki ima zanimivo posledico.

Izrek 1.8 *Vsak izjavni izraz I lahko enakovredno prepišemo samo z uporabo veznikov $\neg, \wedge, \vee$.*

Morda bi se na tem mestu vprašali, zakaj sploh definirati preostale izjavne veznike. Navedimo samo enega od razlogov. Večja množica[3] izjavnih veznikov (ki jih vsaj nekatere znamo učinkovito interpretirati v naravnem jeziku) omogoča krajši zapis izjavnih izrazov.

[3] Množice bomo spoznali v poglavju 3.

V skladu z izrekom 1.8 za množico izjavnih veznikov $\mathcal{N}$ pravimo, da je *poln nabor izjavnih veznikov*, če lahko vsak izjavni izraz I enakovredno prepišemo samo z uporabo izjavnih spremenljivk in logičnih veznikov iz $\mathcal{N}$. Izrek 1.8 torej pravi, da je $\{\neg, \wedge, \vee\}$ poln nabor izjavnih veznikov.

Lema 1.9 *Naj bo $\mathcal{N}$ množica izjavnih veznikov in $\mathcal{Z}$ poln nabor izjavnih veznikov. Če lahko vsak veznik iz nabora $\mathcal{Z}$ izrazimo samo z uporabo veznikov iz $\mathcal{N}$, potem je tudi $\mathcal{N}$ poln nabor izjavnih veznikov.*

Dokaz. Naj bo I poljuben izjavni izraz. Ker je $\mathcal{Z}$ poln nabor izjavnih veznikov, obstaja izjavni izraz $I_{\mathcal{Z}}$, ki je enakovreden I in uporablja izključno veznike iz $\mathcal{Z}$.

Vsak vstop veznika v izrazu $I_{\mathcal{Z}}$ lahko po predpostavki izrazimo z uporabo veznikov iz $\mathcal{N}$ in pridelamo izjavni izraz, enakovreden I, ki vsebuje samo veznike iz $\mathcal{N}$. □

Izrek 1.10 *Nabori $\{\neg, \vee\}$, $\{\neg, \wedge\}$, $\{\neg, \Rightarrow\}$ in $\{\Rightarrow, 0\}$ so polni nabori izjavnih veznikov.*

Dokaz. Dokaze polnosti naborov bomo izdelali v skladu z Lemo 1.9 — pri vsakem od naborov na tapeti bomo izbrali znan polni nabor in izrazili veznike znanega polnega nabora.

Obravnavajmo nekoliko podrobneje nabor $\{\neg, \vee\}$. Pri izbiri polnega nabora nimamo veliko možnosti, $\{\neg, \wedge, \vee\}$ je trenutno edini polni nabor, ki ga poznamo. Za dokaz polnosti nabora $\{\neg, \vee\}$ je potrebno vsakega od veznikov $\neg, \wedge, \vee$ izraziti s pomočjo veznikov $\neg$ in $\vee$. Pri tem je dve tretjini dela že opravljenega, veznika $\neg$ in $\vee$ nastopata v obeh naborih, tako v tistem pod drobnogledom kot v izbranem znanem polnem naboru. V računu

$$p \wedge q \;\sim\; \neg\neg(p \wedge q) \;\sim\; \neg(\neg p \vee \neg q)$$

smo uspeli konjunkcijo $p \wedge q$ zapisati samo z uporabo negacije $\neg$ in disjunkcije $\vee$. Zato je nabor izjavnih veznikov $\{\neg, \vee\}$ poln.

Nabor $\{\neg, \wedge\}$ obdelamo po sorodnem premisleku, lahko bi uporabili celo izrek 1.4 o dualnosti.

Pri naboru $\{\neg, \Rightarrow\}$ za znani polni nabor izberimo $\{\neg, \vee\}$. Disjunkcijo lahko z uporabo negacije in implikacije izrazimo takole

$$p \vee q \;\sim\; \neg\neg p \vee q \;\sim\; \neg p \Rightarrow q.$$

Logično konstanto 0 brez ustrezne matematične abstrakcije težko proglasimo za izjavni veznik. Polnost nabora $\{\Rightarrow, 0\}$ smemo razumeti na način, da lahko vsak izjavni izraz enakovredno zapišemo samo z uporabo implikacij, če smemo poleg izjavnih spremenljivk uporabljati tudi logično konstanto 0.

Dokaz polnosti nabora $\{\Rightarrow, 0\}$ uporabi polni nabor $\{\neg, \Rightarrow\}$. Ker je $\neg p \sim p \Rightarrow 0$, smo uspeli vsakega od veznikov nabora $\{\neg, \Rightarrow\}$ izraziti samo z uporabo veznikov iz $\{\Rightarrow, 0\}$ in dokaz je zaključen. □

Gre morda krajše? Je moč skonstruirati poln nabor izjavih veznikov, ki vsebuje en sam dvomestni izjavni veznik?

Premislimo najprej, da nabor $\{\Rightarrow, \Leftrightarrow\}$ ni poln. Opazimo, da tako veznik $\Rightarrow$ kot veznik $\Leftrightarrow$ ohranjata logično vrednost 1, saj je $1 \Rightarrow 1 \sim 1$ in $1 \Leftrightarrow 1 \sim 1$. Pravimo tudi, da nabor $\{\Rightarrow, \Leftrightarrow\}$ *ohranja logično vrednost* 1.

Izjavni izraz, ki ga sestavimo samo z uporabo izjavnih veznikov $\Rightarrow$ in $\Leftrightarrow$, bo imel v primeru, ko so vse njegove spremenljivke resnične, tudi sam logično vrednost 1. To pomeni, da z uporabo veznikov iz nabora $\{\Rightarrow, \Leftrightarrow\}$ nikakor ne moremo sestaviti izraza, ki bi bil enakovreden izrazu $\neg p$.

Z analognim premislekom pokažemo, da tudi nabori $\{\wedge, \veebar\}$, $\{\wedge, \Rightarrow\}$ in $\{\wedge, \vee\}$ niso polni, saj ohranjajo, po vrsti, logične vrednosti 0, 1 in 0 (slednji dejansko ohranja obe logični vrednosti).

Odtod sledi, da poln nabor, ki vsebuje en sam (dvomestni) logični veznik, nikakor ne sme ohranjati nobene od logičnih vrednosti.

Izrek 1.11 *Nabora* $\{\uparrow\}$ *in* $\{\downarrow\}$ *sta polna nabora izjavnih veznikov in sta edina polna nabora izjavnih veznikov, ki vsebujeta en sam dvomestni izjavni veznik.*

Dokaz. Polnost nabora $\{\uparrow\}$ dokažemo z redukcijo na znani polni nabor $\{\neg, \wedge\}$.

$$\begin{aligned} \neg p &\sim \neg(p \wedge p) \sim p \uparrow p \\ p \wedge q &\sim \neg\neg(p \wedge q) \sim \neg(p \uparrow q) \sim (p \uparrow q) \uparrow (p \uparrow q) \end{aligned}$$

O polnosti nabora $\{\downarrow\}$ lahko premislimo podobno, utemeljitev prepustimo bralcu.

Za dokaz drugega dela trditve obravnavajmo pravilnostno tabelo dvomestnega izjavnega veznika z generično oznako $\square$, za katerega smemo privzeti, da ne ohranja nobene od logičnih vrednosti 0 oziroma 1.

p	q	$p \square q$
0	0	1
0	1	x_1
1	0	x_2
1	1	0

Kateri logični vrednosti x_1, x_2 smemo izbrati? Če je $x_1 = x_2$, potem je veznik $\square$ bodisi enak Shefferjevemu bodisi Peirce-Łukasiewiczevemu vezniku $\uparrow$ oziroma $\downarrow$. Če pa je $x_1 \neq x_2$, potem je izraz $p \square q$ enakovreden bodisi $\neg p$ bodisi $\neg q$. V tem primeru nabor $\{\square\}$ ni poln, saj bo logična vrednost izjavnega izraza sestavljenega izključno z uporabo

veznika $\square$ odvisna le od prve spremenljivke v izrazu (oziroma od zadnje, če je $x_2 = 1$). Nikakor pa ne bi mogli s takšnim veznikom izraziti, denimo, konjunkcije $p \wedge q$. $\square$

1.7 Sklepanje v izjavnem računu

Sklepanje je osnovna naloga matematične logike. Pri izbrani družini izjav-predpostavk nas bo zanimalo, ali smemo logično sklepati na izjavo-zaključek.

V naslednjih sklepih so predpostavke zapisane nad delilno črto, zaključek pa pod njo.

Sklep

$$\frac{\begin{array}{l}\textit{Če dežuje, se Bine ne vozi s kolesom.}\\ \textit{Dežuje.}\end{array}}{\textit{Bine se ne vozi s kolesom.}}$$

se zdi pravilen, sklepa z istimi predpostavkami z drugačnima zaključkoma

$$\frac{\begin{array}{l}\textit{Če dežuje, se Bine ne vozi s kolesom.}\\ \textit{Dežuje.}\end{array}}{\textit{Bine kolesari.}} \qquad \frac{\begin{array}{l}\textit{Če dežuje, se Bine ne vozi s kolesom.}\\ \textit{Dežuje.}\end{array}}{\textit{Sonce sije ali pa Bine vozi avto.}}$$

pa se zdita napačna. Zakaj?

Pravilnost sklepa definirajmo raje v domeni izjavnih izrazov. Zaporedje izjavnih izrazov

$$A_1, \ldots, A_k, B$$

je *pravilen sklep* s *predpostavkami* $A_1, \ldots, A_k$ in *zaključkom* B, če je pri vseh naborih logičnih vrednosti spremenljivk (ki nastopajo v predpostavkah in zaključku), pri katerih so vse predpostavke resnične, resničen tudi zaključek.

V tem primeru pišemo

$$A_1, \ldots, A_k \models B, \tag{1.3}$$

in preberemo, da zaključek B *logično sledi* iz predpostavk $A_1, \ldots, A_k$.

Preberimo definicijo pravilnega sklepa še drugače. V resnici nas zanimajo samo tisti nabori logičnih vrednosti spremenljivk, pri katerih so vse predpostavke resnične. In v vseh takšnih naborih mora biti resničen tudi zaključek sklepa. Nabor, pri katerem niso izpolnjene vse predpostavke, na pravilnost sklepa ne vpliva. Logična vrednost zaključka je v takšnem naboru lahko poljubna.

Z naslednjim poimenovanjem enostavnih izjav

d ... *Dežuje.*
k ... *Bine se vozi s kolesom.*
s ... *Sonce sije.*
a ... *Bine vozi avto.*

lahko začetne zglede sklepo prepišemo v

$$\frac{d \Rightarrow \neg k \quad d}{\neg k} \qquad \frac{d \Rightarrow \neg k \quad d}{k} \qquad \frac{d \Rightarrow \neg k \quad d}{s \vee a} \tag{1.4}$$

S pravilnostno tabelo se lahko prepričamo, da je prvi izmed sklepov (1.4) pravilen, s krepko pisavo je označen edini nabor logičnih vrednosti spremenljivk, pri katerem so resnične vse predpostavke. V tem primeru je resničen tudi zaključek.

d	k	$d \Rightarrow \neg k$	d	$\neg k$
0	0	1	0	1
0	1	1	0	0
1	**0**	**1**	**1**	**1**
1	1	0	1	0

Preostala dva sklepa (1.4) sta napačna, saj lahko najdemo *protiprimer* — to je nabor logičnih vrednosti, pri katerem so vse predpostavke resnične, zaključek pa neresničen. Z izbiro nabora logičnih vrednosti, v katerem je spremenljivka d resnična, vse ostale spremenljivke k, s in a pa imajo logično vrednost 0, sta v obeh primerih predpostavki resnični, zaključka pa napačna.

Tudi sklep

$$\frac{\begin{array}{l} r \vee \neg s \\ p \Rightarrow q \\ p \vee s \end{array}}{r \vee \neg q} \tag{1.5}$$

ni pravilen. Protiprimer je nabor logičnih vrednosti $p \sim 1$, $q \sim 1$, $r \sim 0$ in $s \sim 0$. Z direktnim računom

$$\frac{\begin{array}{l} r \vee \neg s \ \sim\ 0 \vee 1 \ \sim\ 1 \\ p \Rightarrow q \ \sim\ 1 \Rightarrow 1 \ \sim\ 1 \\ p \vee s \ \sim\ 1 \vee 0 \ \sim\ 1 \end{array}}{r \vee \neg q \ \sim\ 0 \vee 0 \ \sim\ 0}$$

pridelamo, da so pri tem naboru vse predpostavke resnične, zaključek pa ne.

Pravilnost sklepa lahko, če se ne oziramo na dolžino dokaza, preverimo z uporabo pravilnostnih tabel vseh predpostavk in zaključka. Ni presenetljivo, da lahko pravilnost sklepa prepišemo v pravilnost enega samega izjavnega izraza.

Izrek 1.12 *$A_1, A_2, \ldots, A_n \models B$ natanko tedaj, ko je izjavni izraz $(A_1 \wedge A_2 \wedge \ldots \wedge A_n) \Rightarrow B$ tavtologija.*

Dokaz. Izjavni izraz $(A_1 \wedge A_2 \wedge \ldots \wedge A_n) \Rightarrow B$ je tavtologija natanko tedaj, ko je pri vsakem naboru logičnih vrednosti resničen. To pomeni, da je v vsakem naboru logičnih vrednosti, pri katerih je resničen antecedens $A_1 \wedge A_2 \wedge \ldots \wedge A_n$, resničen tudi konsekvens B.

Enakovredno, v vsakem naboru, pri katerem so resnični vsi izrazi-predpostavke $A_1, \ldots, A_k$, je resničen tudi zaključek B. □

Navedimo še nekaj lastnosti sklepov.

Izrek 1.13

(S1) *Če je $B \sim C$, potem $A_1, \ldots, A_k \models B$ natanko tedaj, ko $A_1, \ldots, A_k \models C$.*

(S2) $A_1, \ldots, A_k \models 1$

(S3) $0, A_1, \ldots, A_k \models B$

(S4) *Za vsak $i \in \{1, \ldots, k\}$ velja: $A_1, A_2, \ldots, A_k \models A_i$*

(S5) $A_1, \ldots, A_k \models B$ *natanko tedaj, ko* $A_1, A_2, \ldots, A_k, 1 \models B$

Dokaz. Za dokaz (S1) upoštevamo, da imata B in C v vseh naborih iste logične vrednosti. Zato zamenjava izraza B z izrazom C ne vpliva na spremembo pravilnosti sklepa.

Pri (S2) in (S3) opazimo, da je bodisi zaključek pravilen pri vsakem naboru bodisi ne obstaja nabor, pri katerem so vse predpostavke pravilne. Podoben premislek ukani (S4) — če so pravilne vse predpostavke, potem je pravilna vsaka izmed njih.

Privzemanje dodatne predpostavke v obliki tavtologije ne vpliva na množico naborov, pri katerih so vse predpostavke pravilne. Zato velja tudi (S5). □

1.7.1 Pravila sklepanja in dokaz pravilnosti sklepa

Dokaz nepravilnosti sklepa smo spravili pod streho, dovolj je poiskati protiprimer, en sam ustrezen nabor logičnih vrednosti. Kako pa pokažemo pravilnost sklepa? Zapis pravilnostne tabele in preverjanje vseh naborov je časovno potraten postopek. Precej rajši bi imeli kratko izpeljavo, v kateri bi izvajali relativno enostavne, majhne korake proti cilju.

Majhne, enostavne sklepe, ki jih bomo potrebovali za dokazovanje pravilnosti sklepov, imenujemo *pravila sklepanja*.

modus ponens	**MP**	$A, A \Rightarrow B$	$\models B$
modus tollens	**MT**	$A \Rightarrow B, \neg B$	$\models \neg A$
hipotetični silogizem	**HS**	$A \Rightarrow B, B \Rightarrow C$	$\models A \Rightarrow C$
disjunktivni silogizem	**DS**	$A \vee B, \neg A$	$\models B$
združitev	**Zd**	A, B	$\models A \wedge B$
poenostavitev	**Po**	$A \wedge B$	$\models A$
pridružitev	**Pd**	A	$\models A \vee B$

Tabela 1.6: Pravila sklepanja

Izrek 1.14 *Pravila sklepanja, predstavljena v tabeli 1.6, so pravilni sklepi.*

Dokaz. Dokažimo pravilnost hipotetičnega silogizma z uporabo pravilnostne tabele.

A	B	C	$A \Rightarrow B$	$B \Rightarrow C$	$A \Rightarrow C$
0	**0**	**0**	**1**	**1**	**1**
0	**0**	**1**	**1**	**1**	**1**
0	1	0	1	0	1
0	**1**	**1**	**1**	**1**	**1**
1	0	0	0	1	0
1	0	1	0	1	1
1	1	0	1	0	0
1	**1**	**1**	**1**	**1**	**1**

Nabori logičnih vrednosti izrazov A, B, C, pri katerih so vse predpostavke resnične, so zapisani polkrepko. V vseh omenjenih naborih je pravilen tudi zaključek $A \Rightarrow C$, hipotetični silogizem je torej pravilen sklep.

Modus tollens ukanimo z uporabo izreka 1.12. Pokazati je potrebno, da je izraz

$$(A \Rightarrow B) \wedge \neg B \Rightarrow \neg A$$

tavtologija. Računajmo:

$$\begin{aligned}
(A \Rightarrow B) \wedge \neg B \Rightarrow \neg A &\sim \neg((A \Rightarrow B) \wedge \neg B) \vee \neg A \\
&\sim \neg((\neg A \vee B) \wedge \neg B) \vee \neg A \\
&\sim (\neg(\neg A \vee B) \vee B) \vee \neg A \\
&\sim (A \wedge \neg B) \vee B \vee \neg A \\
&\sim (A \vee B \vee \neg A) \wedge (\neg B \vee B \vee \neg A) \\
&\sim (A \vee \neg A \vee B) \wedge (\neg B \vee B \vee \neg A) \\
&\sim (1 \vee B) \wedge (1 \vee \neg A) \ \sim \ 1 \wedge 1 \ \sim \ 1
\end{aligned}$$

Dokaze pravilnosti preostalih pravil sklepanja prepustimo bralcu. □

Pravila sklepanja modus ponens, modus tollens, disjunktivni in hipotetični silogizem so znana že iz antičnih časov. Pravila sklepanja poenostavitev, združitev in pridružitev so po drugi strani precej lažje razumljiva in zelo uporabna pri manipulaciji izrazov v dokazovanju pravilnosti sklepa.

Dokaz pravilnosti sklepa

$$A_1, \ldots, A_k \models B$$

je zaporedje izjavnih izrazov

$$C_1, C_2, \ldots, C_m \tag{1.6}$$

v katerem je $C_m = B$, sicer pa za vsak posamezen člen C_i velja

(DP1) C_i je predpostavka ali

(DP2) C_i je tavtologija ali

(DP3) C_i je enakovreden enemu od prejšnjih členov zaporedja (1.6) ali

(DP4) C_i logično sledi iz prejšnjih členov zaporedja (1.6) po enem od pravil sklepanja.

Izrek 1.15 *Denimo, da je zaporedje*

$$C_1, C_2, \ldots, C_m$$

dokaz pravilnosti sklepa

$$A_1, \ldots, A_k \models B.$$

Potem so v vsakem naboru, pri katerem so resnične vse predpostavke $A_1, \ldots, A_k$, resnični tudi vsi izrazi C_j, $j \in \{1, \ldots, m\}$, iz zaporedja.

Dokaz. Izberimo nabor logičnih vrednosti spremenljivk, pri katerem so vse predpostavke resnične. Izraz C_1 je predpostavka ali tavtologija, zato je v tem naboru resničen.

Opazujmo C_j in induktivno privzemimo, da so vsi izrazi $C_1, \ldots, C_{j-1}$ v tem naboru resnični. Če je C_j predpostavka, tavtologija ali enakovreden enemu od prejšnjih členov zaporedja, potem je C_j pri tem naboru zagotovo resničen.

Preostane nam še obravnava možnosti, ko C_j logično sledi iz prejšnjih členov po katerem od pravil sklepanja. Pravila sklepanja so po izreku 1.14 pravilni sklepi. Če je C_j zaključek takšnega pravilnega sklepa, mora biti pravilen v primeru pravilnih predpostavk. Le-te so vsebovane med členi $C_1, \ldots, C_{j-1}$ in zato po indukcijski predpostavki pravilne. □

Izrek 1.15 trdi tudi, da nepravilnih sklepov ne moremo dokazati. Protiprimer nepravilnega sklepa je namreč nabor logičnih vrednosti, pri katerem so predpostavke resnične,

zaključek pa ne. Takšnega zaključka, ki logično ne sledi iz predpostavk, torej ne moremo pridelati kot enega od členov dokaza pravilnosti.

Kaj pa pravilni sklepi? Ali je za vsak pravilen sklep moč poiskati dokaz njegove pravilnosti? Izkaže se, da je temu res tako. Če $A_1, \ldots, A_k \models B$, je po izreku 1.12 izraz $(A_1 \wedge \ldots \wedge A_k) \Rightarrow B$ tavtologija. S $(k-1)$-kratno uporabo pravila združitve iz predpostavk pridelamo konjunkcijo $A_1 \wedge \ldots \wedge A_k$, iz te in iz gornje tavtologije pa po pravilu modus ponens dobimo zaključek B.

Izdelajmo za začetek dokaz pravilnosti naslednjega sklepa.

$$\begin{array}{l} p \Rightarrow q \\ p \vee s \\ \neg s \wedge \neg r \\ \hline q \wedge \neg r \end{array} \tag{1.7}$$

Dokaz pravilnosti bomo zapisali v oštevilčenih vrsticah. Zaporedna številka vrstice pomeni mesto posameznega izjavnega izraza v dokazu pravilnosti sklepa. Obenem bomo navedli tudi utemeljitev, zakaj posamezni izjavni izraz lahko zapišemo v dokaz pravilnosti:

1.	$p \Rightarrow q$	predpostavka
2.	$p \vee s$	predpostavka
3.	$\neg s \wedge \neg r$	predpostavka
4.	$\neg s$	Po(3)
5.	p	DS(4,2)
6.	q	MP(5,1)
7.	$\neg r$	Po(3)
8.	$q \wedge \neg r$	Zd(6,7)

V dokazu pravilnosti sklepa lahko v prvi vrstici po (DP1) zapišemo predpostavko. To bomo v dokazovanju tudi tipično storili, namreč, na začetku prepisali vse predpostavke.

1.8 Pomožni sklepi

Poskusimo izdelati tudi dokaz naslednjega pravilnega sklepa.

$$\begin{array}{l} \neg(p \wedge q) \\ r \Rightarrow q \\ r \vee s \\ \hline p \Rightarrow s \end{array} \tag{1.8}$$

Začnimo z zapisom predpostavk.

1.	$\neg(p \wedge q)$	predpostavka

2.	$r \Rightarrow q$	predpostavka
3.	$r \vee s$	predpostavka

Na tem mestu se znajdemo v težavah. Ni čisto jasno, kateri izraz zapisati kot naslednjega v dokaz pravilnosti sklepa. Sicer lahko izraza vedno *združimo* s konjunkcijo ali izrazu *pridružimo* katerikoli novi izraz, nismo pa prepričani, da omenjena koraka vodita k uspehu. Preostalih pravil sklepanja pa v tem trenutku ne moremo uporabiti.

Poskusimo s preoblikovanjem izrazov po enakovrednosti.

4.	$\neg p \vee \neg q$	$\sim$ (1)
5.	$p \Rightarrow \neg q$	$\sim$ (4)
6.	$\neg q \Rightarrow \neg r$	$\sim$ (2)
7.	$p \Rightarrow \neg r$	HS(5,6)
8.	$\neg\neg r \vee s$	$\sim$ (3)
9.	$\neg r \Rightarrow s$	$\sim$ (8)
10.	$p \Rightarrow s$	HS(7,9)

Uspelo nam je šele z večkratno uporabo enakovrednosti, uporabili pa smo zgolj dve pravili sklepanja.

Pomožni sklepi so orodja, kot nekakšni pomožni računi, ki nam bodo omogočili, da v dokazovanju pravilnosti sklepa čim pogosteje uporabljamo pravila sklepanja. S tem se bomo po eni strani izognili pretirani uporabi enakovrednosti (posamezen izraz lahko z uporabo različnih zakonov izjavnega računa enakovredno preoblikujemo v precej različnih izjavnih izrazov — pri tem pa ne vemo vnaprej, katerega bi potrebovali), poleg tega pa bomo z uporabo pravil sklepanja ohranjali izraze v zaporedju dokaza kar se da kratke.

Pogojni sklep

Prvi pomožni sklep je *pogojni sklep*, uporabljamo ga lahko takrat, ko ima zaključek obliko implikacije. Srž pogojnega sklepa je skrita v naslednjem izreku.

Izrek 1.16 (pogojni sklep PS) $A_1, \ldots, A_k \models B \Rightarrow C$ *natanko tedaj, ko* $A_1, \ldots, A_k, B \models C$.

Dokaz. Izrek 1.12 enači pravilnost sklepov

$$A_1, \ldots, A_k \models B \Rightarrow C$$

oziroma

$$A_1, \ldots, A_k, B \models C$$

z dejstvom, da sta izjavna izraza

$$(A_1 \wedge \ldots \wedge A_k) \Rightarrow (B \Rightarrow C) \tag{1.9}$$

oziroma

$$(A_1 \wedge \ldots \wedge A_k \wedge B) \Rightarrow C \tag{1.10}$$

tavtologiji. Zato je dovolj premisliti, da sta izraza (1.9) in (1.10) enakovredna. Računajmo:

$$\begin{aligned}(A_1 \wedge \ldots \wedge A_k) \Rightarrow (B \Rightarrow C) &\sim \neg(A_1 \wedge \ldots \wedge A_k) \vee (\neg B \vee C)\\ &\sim \neg((A_1 \wedge \ldots \wedge A_k) \wedge B) \vee C\\ &\sim (A_1 \wedge \ldots \wedge A_k \wedge B) \Rightarrow C\end{aligned}$$

Dokaz je s tem pri koncu. □

Uporaba pogojnega sklepa spremeni množico predpostavk — v množico prvotnih predpostavk vrine dodatno predpostavko, pravimo ji predpostavka pogojnega sklepa, ki je enaka antecedensu implikacije v zaključku. Obenem spremeni tudi zaključek — nadomesti ga s konsekvensom originalnega zaključka.

Z uporabo dodatnih predpostavk nam pogojni sklep poveča možnost uporabe pravil sklepanja v dokazu. Ravno tako nam poenostavi zaključek, njegova nova globina je strogo manjša kot originalna.

Mehanizem zapisa dokaza z uporabo pogojnega sklepa si bomo ogledali na primeru pravilnega sklepa (1.8).

1.	$\neg(p \wedge q)$	predpostavka
2.	$r \Rightarrow q$	predpostavka
3.	$r \vee s$	predpostavka
4.	$\neg p \vee \neg q$	$\sim$ (1)
5.1.	p	predpostavka PS
5.2.	$\neg q$	DS(4,5.1)
5.3.	$\neg r$	MT(2,5.2)
5.4.	s	DS(3,5.3)
5.	$p \Rightarrow s$	PS(5.1,5.4)

Uporaba pogojnega sklepa se skriva v zamaknjenih vrsticah med 5.1 in 5.4. V vrstici 5.1 uvedemo dodatno predpostavko pogojnega sklepa, sam pogojni sklep pa je strnjen v (sestavljeni) vrstici 5.

Zakaj je potrebno zamikanje vrstic? Mika nas, da bi dokaz pravilnosti sklepa nadaljevali s še eno vrstico

6.	s	MP(5.1,5)	$\longleftarrow$ Pazi, napaka!!

vendar tega dejansko ne smemo storiti. Sklep

$$\begin{array}{l}\neg(p \wedge q)\\ r \Rightarrow q\\ r \vee s\\ \hline s\end{array} \tag{1.11}$$

je namreč nepravilen, protiprimer je naslednji nabor logičnih vrednosti

p	q	r	s	$\neg(p \wedge q)$	$r \Rightarrow q$	$r \vee s$	s
0	1	1	0	1	1	1	0

Težava je v zamaknjenih vrsticah med 5.1 in 5.4. Le-teh po zaključku pogojnega sklepa (vrstici 5) ne smemo več uporabljati.

Uporaba pogojnega sklepa je možna, če imamo v zaključku implikacijo. Tudi za dokaz disjunkcije smemo uporabiti pogojni sklep, saj lahko zaključek $A \vee B$ nadomestimo z implikacijo $\neg A \Rightarrow B$ ali celo z alternativo $\neg B \Rightarrow A$.

Sklep s protislovjem

Naslednji pomožni sklep je *sklep s protislovjem* (lat. *reductio ad absurdum*, uporabljamo tudi kratico RA), ki ga lahko uporabljamo ne glede na obliko zaključka. Mehanizem sklepa s protislovjem opisuje naslednji rezultat.

Izrek 1.17 (sklep s protislovjem RA) *$A_1, A_2, \ldots, A_k \models B$ natanko tedaj, ko $A_1, A_2, \ldots, A_k, \neg B \models 0$.*

Dokaz. Postopamo kot v dokazu izreka 1.16 in dokazujemo enakovrednost izrazov

$$(A_1 \wedge \ldots \wedge A_k) \Rightarrow B \tag{1.12}$$

in

$$(A_1 \wedge \ldots \wedge A_k \wedge \neg B) \Rightarrow 0 \tag{1.13}$$

z naslednjim računom.

$$\begin{aligned}(A_1 \wedge \ldots \wedge A_k \wedge \neg B) \Rightarrow 0 &\sim \neg(A_1 \wedge \ldots \wedge A_k \wedge \neg B) \vee 0 \\ &\sim \neg(A_1 \wedge \ldots \wedge A_k \wedge \neg B) \\ &\sim \neg(A_1 \wedge \ldots \wedge A_k) \vee B \\ &\sim (A_1 \wedge \ldots \wedge A_k) \Rightarrow B\end{aligned}$$

Ker sta izraza (1.12) in (1.13) enakovredna, sta hkrati oba tavtologiji. Torej sta tudi sklepa iz formulacije izreka hkrati pravilna. □

Dokažimo pravilnost sklepa (1.8) še z uporabo sklepa s protislovjem.

1.	$\neg(p \wedge q)$	predpostavka
2.	$r \Rightarrow q$	predpostavka
3.	$r \vee s$	predpostavka

4.	$\neg p \vee \neg q$	$\sim$ (1)
5.1.	$\neg(p \Rightarrow s)$	predpostavka RA
5.2.	$p \wedge \neg s$	$\sim$ (5.1)
5.3.	p	Po(5.2)
5.4.	$\neg s$	Po(5.2)
5.5.	$\neg q$	DS(4,5.3)
5.6.	$\neg r$	MT(2,5.5)
5.7.	s	DS(3,5.6)
5.8.	$s \wedge \neg s$	Zd(5.7,5.4)
5.9.	0	$\sim$ (5.9)
5	$p \Rightarrow s$	RA(5.1,5.8)

Na hitro primerjajmo pogojni sklep in sklep s protislovjem v primeru, ko ima zaključek sklepa obliko implikacije $B \Rightarrow C$. Sklep s protislovjem med predpostavke doda negacijo originalnega zaključka $\neg(B \Rightarrow C) \sim B \wedge \neg C$. Če konjunkcijo $B \wedge \neg C$ z uporabo poenostavitve zapišemo kot par dodatnih predpostavk $B, \neg C$, lahko uporabo sklepa s protislovjem v tem primeru razumemo na naslednji način:

Pogojni sklep med predpostavke doda eno dodatno predpostavko, sklep s protislovjem pa kar dve. Zaključek pogojnega sklepa je enostavnejši v primerjavi z originalnim zaključkom, sklep s protislovjem gre še korak naprej, njegov zaključek je kar logična konstanta.

Čez prst bi lahko dejali, da je sklep s protislovjem še za eno stopnjo močnejše orodje za dokazovanje pravilnosti sklepa v primerjavi s pogojnim sklepom.

Analiza primerov

Za konec navedimo še zadnji pomožni sklep, *analizo primerov*. Analizo primerov uporabljamo, ko ima katera izmed predpostavk obliko disjunkcije. Ker lahko tavtologijo 1 *vedno* vtaknemo med predpostavke in jo nadomestimo z enakovrednim izrazom, denimo, s $p \vee \neg p$, imamo, vsaj v principu, med predpostavkami vedno lahko takšno, ki ustreza zahtevi.

Način delovanja analize primerov opisuje naslednji izrek.

Izrek 1.18 (analiza primerov — AP) *$A_1, \ldots, A_n, B_1 \vee B_2 \models C$ natanko tedaj, ko $A_1, \ldots, A_k, B_1 \models C$ in $A_1, \ldots, A_k, B_2 \models C$.*

Dokaz. Znova se skličemo na izrek 1.12. Pokazati je potrebno, da je izraz

$$(A_1 \wedge \ldots \wedge A_k \wedge (B_1 \vee B_2)) \Rightarrow C \tag{1.14}$$

tavtologija natanko tedaj, ko sta tavtologiji oba izraza

$$(A_1 \wedge \ldots \wedge A_k \wedge B_1) \Rightarrow C \quad \text{in} \quad (A_1 \wedge \ldots \wedge A_k \wedge B_2) \Rightarrow C. \tag{1.15}$$

Ponovno računajmo.

$$
\begin{aligned}
(A_1 \wedge \ldots \wedge A_k \wedge (B_1 \vee B_2)) \Rightarrow C \ &\sim \\
&\sim \ \neg(A_1 \wedge \ldots \wedge A_k \wedge (B_1 \vee B_2)) \vee C \\
&\sim \ \neg(A_1 \wedge \ldots \wedge A_k) \vee \neg(B_1 \vee B_2) \vee C \\
&\sim \ \neg(A_1 \wedge \ldots \wedge A_k) \vee (\neg B_1 \wedge \neg B_2) \vee C \\
&\sim \ (\neg(A_1 \wedge \ldots \wedge A_k) \vee \neg B_1 \vee C) \wedge (\neg(A_1 \wedge \ldots \wedge A_k) \vee \neg B_2 \vee C) \\
&\sim \ (\neg(A_1 \wedge \ldots \wedge A_k \wedge B_1) \vee C) \wedge (\neg(A_1 \wedge \ldots \wedge A_k \wedge B_2) \vee C) \\
&\sim \ ((A_1 \wedge \ldots \wedge A_k \wedge B_1) \Rightarrow C) \wedge ((A_1 \wedge \ldots \wedge A_k \wedge B_2) \Rightarrow C)
\end{aligned}
$$

Za konec upoštevajmo samo še dejstvo, da je konjunkcija tavtologija natanko tedaj, ko sta tavtologiji oba njena člena. □

Zapišimo še dokaz pravilnosti sklepa (1.8) z uporabo analize primerov, številčenje je v tem primeru nekoliko bolj zapleteno.

1.	$\neg(p \wedge q)$	predpostavka
2.	$r \Rightarrow q$	predpostavka
3.	$r \vee s$	predpostavka
4.	$\neg p \vee \neg q$	$\sim$ (1)
5.1.1	r	predpostavka AP_1
5.1.2	q	MP(5.1.1,2)
5.1.3	$\neg p$	DS(5.1.2,4)
5.1.4	$\neg p \vee s$	Pr(5.1.3)
5.2.1	s	predpostavka AP_2
5.2.2.	$\neg p \vee s$	Pr(5.2.1)
5.	$\neg p \vee s$	AP(3,5.1,5.2)
6.	$p \Rightarrow s$	$\sim$ (5)

Pri tem številčenje AP(3,5.1,5.2) pomeni, da smo analizo primerov AP uporabili na disjunkciji v 3. vrstici dokaza, obe možnosti analize primerov pa nastopata v vejah 5.1 in 5.2.

Analizo primerov smemo seveda uporabljati tudi na disjunkcijah z več kot le dvema členoma.

Analiza primerov glede na disjunkcijo $p \vee \neg p$ — takšna disjunkcija je enakovredna tavtologiji in jo lahko vedno podtaknemo med predpostavke — loči primera, ko je p lažen oziroma p resničen. V vsakem posameznem primeru bi lahko z analizo primerov obravnavali obe logični vrednosti druge spremenljivke, nato tretje itn. Uporaba analize primerov glede na logične vrednosti vseh nastopajočih spremenljivk pa pripelje dokazovanje pravilnosti sklepa ponovno v območje pravilnostnih tabel — le-te pa, kot vemo, so lahko zelo obsežne.

Poglavje 2

Predikatni račun

2.1 Zakaj predikatni račun

V prvem poglavju smo se srečali z izjavnim računom. Zgodbo smo pripeljali dovolj daleč, da znamo povedati, kateri sklepi so pravilni in zakaj. V zaporedju izjavnih izrazov znamo dokazati, zakaj zaključek logično sledi iz predpostavk, oziroma poiskati protiprimer, če temu ni tako.

Izkaže se, da je zgodba (matematične) logike precej bolj zapletena, kot se nam zdi po prebiranju prvega poglavja.

Oglejmo si naslednji sklep.

$$\begin{array}{l} \textit{Vse krave dajejo mleko.} \\ \textit{Tarzan ne daje mleka.} \\ \hline \textit{Tarzan ni krava.} \end{array} \tag{2.1}$$

S stališča izjavnega računa so vse izjave, tako predpostavke kot zaključek, v sklepu (2.1) enostavne. Če jih želimo zapisati kot izjavne izraze, nimamo nobene boljše možnosti, kot da jih po vrsti zapišemo z izjavnimi spremenljivkami p, q, r.

To je seveda daleč od dejanskega stanja. Ime *"Tarzan"* se pojavi tako v eni od predpostavk kot tudi v zaključku sklepa, pojem *"je krava"* ravno tako. *"Mleko"* je izraz, ki povezuje predpostavki. Izjave, predpostavke in zaključek, so med seboj povezane, problem je v izjavnem računu. Izjavni račun takšnih povezav ne zmore zaznati in izraziti.

Predikatni račun je nadgradnja izjavnega računa, opisati bomo mogli tudi notranjo zgradbo posamezne enostavne izjave in s pomočjo notranjih struktur upravičiti pravilnost sklepa (2.1).

Predikati

Stavek *"x daje mleko"*, ni izjava. Njegova logična vrednost je namreč odvisna od izbire vrednosti spremenljivke x. Če na njeno mesto v omenjeni stavek vstavimo *"Psa Flokija"*, dobimo lažno izjavo, v primeru, ko vstavimo *"Lisko iz sosedove črede molznic"*, pa resnično izjavo.

Dejstvo, da neki osebek *"daje mleko"*, smemo razumeti kot njegovo lastnost. Če se pogovarjamo o *živih bitjih in literarnih in filmskih junakih*, potem imajo nekateri izmed njih lastnost *"dajati mleko"*, nekateri izmed njih pa te lastnosti nimajo.

Predikati so logične preslikave[1], v katere vstavljamo konkretne elemente področja pogovora. Če za področje pogovora izberemo *ljudi*, *števila* in *živalske vrste*, nam predikat S lahko pomeni

$$\begin{array}{lll} S(x) & \ldots & x \text{ ima sestro.} \\ S(x) & \ldots & x \text{ je sodo število.} \\ S(x) & \ldots & x \text{ je sesalec.} \end{array}$$

Predikati smejo imeti tudi več argumentov. V istih področjih pogovora lahko dvomestni predikat D pomeni

$$\begin{array}{lll} D(x,y) & \ldots & x \text{ je daljni sorodnik od } y. \\ D(x,y) & \ldots & x \text{ je delitelj števila } y. \\ D(x,y) & \ldots & x \text{ živi dlje kot } y. \end{array}$$

Seveda so lahko predikati tudi večmestni. Tako bi z $V(x,y,z)$ lahko opisali stavek *"x je vsota števil y in z"* v področju pogovora števil, ali pa stavek *"x in y sta roditelja z"* predstavili kot $R(x,y,z)$.

Potegnimo črto: Pri izbranem področju pogovora bodo enomestni predikati opisovali lastnosti elementov področja pogovora, le-ti takšno lastnost lahko imajo ali pa ne. Dvomestni predikati opisujejo zveze (relacije) med pari elementov področja pogovora, k-mestni predikati pa zveze med k-terkami elementov področja pogovora.

Univerzalni kvantifikator

Predikat P definirajmo z opisom

$$P(x) \quad \ldots \quad x \text{ je praštevilo.}$$

V tem primeru je

$$P(3) \wedge P(5) \wedge P(7) \wedge P(11)$$

konjunkcija izjav, ki jo preberemo kot *"3, 5, 7 in 11 so praštevila."* Če se omejimo na družino števil $\mathcal{D} = \{3,5,7,11\}$ bi smeli zapisati tudi

[1] Preslikave nas čakajo v Poglavju 5.

Vsako število iz $\mathcal{D}$ ima lastnost P.

Če na tem mestu pozabimo na področje pogovora in opis predikata P, lahko stavek

Za vsak x velja $P(x)$.

simbolično prepišemo kot

$$\forall x\ P(x) \tag{2.2}$$

Pri tem (2.2) ni izjava, njena logična vrednost je namreč odvisna tako od izbire področja pogovora kot od pomena predikata P.

Simbol $\forall$ imenujemo *univerzalni kvantifikator*. Z univerzalnim kvantifikatorjem lahko izražamo posplošene konjunkcije.

Eksistenčni kvantifikator

Zgodbo lahko recikliramo z uporabo disjunkcije. Predikat Q definirajmo z opisom

$$Q(x) \quad \ldots \quad x \text{ je popoln kvadrat.}$$

V tem primeru je

$$Q(6) \vee Q(7) \vee Q(8) \vee Q(9)$$

disjunkcija izjav, ki jo lahko preberemo kot

Vsaj eno od števil 6, 7, 8, 9 je popoln kvadrat.

ali

Med števili 6, 7, 8, 9 obstaja takšno, ki je popoln kvadrat.

Če na tem mestu pozabimo na področje pogovora in opis predikata Q, lahko stavek

Obstaja x, za katerega velja $Q(x)$.

simbolično prepišemo kot

$$\exists x\ Q(x) \tag{2.3}$$

Tudi (2.3) ni (več) izjava, saj je njena logična vrednost odvisna tako od izbire področja pogovora kot od pomena predikata Q.

Simbol $\exists$ imenujemo *eksistenčni kvantifikator*. Z eksistenčnim kvantifikatorjem lahko izražamo posplošene disjunkcije.

2.2 Izjavne formule

Gradnjo izjavnih izrazov smo začeli z izjavnimi spremenljivkami in izjavnimi konstantami, ki so bile naši najenostavnejši izjavni izrazi. V primeru predikatnega računa bomo naše izraze imenovali *izjavne formule*, s samo izbiro imena jih bomo uspeli ločiti od izjavnih izrazov.

Izjavne formule bodo ravno tako kot izjavni izrazi definirane induktivno. Razlika pa nastopi že pri definiciji najenostavnejših izjavnih formul.

Jezik predikatnega računa vsebuje

(1) *izjavne spremenljivke* $x, y, z, \ldots$

(2) *izjavne konstante* $a, b, c, \ldots$

(3) *predikate* $P, Q, R, \ldots$

(4) izjavne veznike $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, \ldots$

(5) kvantifikatorja $\forall$ in $\exists$ ter

(6) oklepaja (in).

Mestnost posameznega predikata (predikatne črke) z njeno izbiro ne bo določena. Dopuščali bomo tudi, da v isti formuli ista črka P označuje predikata z različnima mestnostima. V splošnem bomo za izjavne konstante res izbirali male črke z začetka abecede, v posebnih primerih jih bomo prilagodili trenutnemu področju pogovora. Tako bomo pri pogovoru o naravnih številih za konstante raje uporabljali konstante $0, 1, 2, 3, \ldots$ kot $a, b, c, d, \ldots$.

Izjavnim spremenljivkam in izjavnim konstantam pravimo z eno besedo *termi*, a, b, x in y so primeri termov. *Atome* (ali *atomarne formule*) pridelamo, če terme vstavljamo v predikate. Zgledi atomov so

$$P(x),\ P(a),\ Q(x, y),\ Q(y, b),\ R(a, x, x).$$

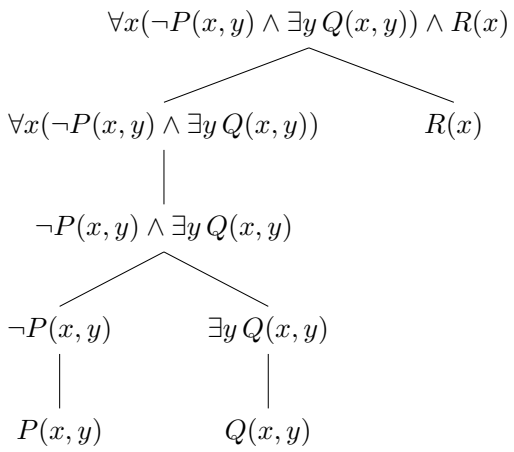
Izjavne formule definiramo induktivno. Pravilo (F1) poskrbi za začetne izjavne formule, s pravilom (F2) obstoječe formule induktivno sestavljamo v nove izjavne formule.

(F1) Atomi so izjavne formule.

(F2) Če sta V in W izjavni formuli in je x izjavna spremenljivka, potem so tudi

$$(\neg W), \quad (W \wedge V), \quad (W \vee V), \quad (W \Rightarrow V), \ \ldots \quad \text{in}$$
$$(\forall x\ W), \quad (\exists x\ W)$$

izjavne formule.



Slika 2.1: Konstrukcijsko drevo izjavne formule $\forall x(\neg P(x,y) \wedge \exists y\, Q(x,y)) \wedge R(x)$.

Na mestu je nekaj komentarjev. Pravilo (F2) poskrbi za konstrukcijo novih izjavnih formul iz obstoječih na dva načina, bodisi z uporabo kvantifikatorjev bodisi z uporabo izjavnih veznikov. Pri tem smemo uporabljati vse izjavne veznike, ne samo tistih, ki smo jih definirali v prejšnjem poglavju.

Predikatni račun smo najavili kot nadgradnjo izjavnega računa, toda kam so se izgubile logične konstante? Logični konstanti pa tudi logične spremenljivke si v predikatnem računu lahko predstavljamo kot 0-mestne predikate — takšne, katerih logična vrednost ni odvisna od izbire argumentov.

2.2.1 Konstrukcijsko drevo

Konstrukcijsko drevo izjavne formule $\forall x(\neg P(x,y) \wedge \exists y\, Q(x,y)) \wedge R(x)$ je predstavljeno na sliki 2.1.

Pri tem smo se v izogib preveliki oklepajski gneči dogovorili za prednost izjavnih veznikov in kvantifikatorjev. Za izjavne veznike uporabljamo dogovor o prednosti iz prejšnjega poglavja, za kvantifikatorja $\forall$ in $\exists$ pa se dogovorimo, da vežeta tako močno kot negacija.

Izjavni formuli W in W', ki se razlikujeta samo v nekaterih parih nepotrebnih oklepajev, smemo *enačiti*. Zapišimo nekaj parov enakih izjavnih formul, ki se razlikujejo samo v nepotrebnih oklepajih.

$$\begin{aligned} \forall x\, \exists y\, \neg P(x,y) &= \forall x(\exists y(\neg P(x,y))) \\ \forall x\, P(x) \wedge Q(x) &= (\forall x\, P(x)) \wedge Q(x) \\ \exists y\, \neg R(x,y) \Rightarrow P(x,y) &= (\exists y(\neg R(x,y))) \Rightarrow P(x,y) \end{aligned}$$

Tudi v domeni izjavnih formul uporabljamo pojme *nastopanja*, *globine* in *dolžine* na isti

način kot v izjavnem računu. Tako formuli $P(x, y)$ in $\exists y\, Q(x, y)$ nastopata v izjavni formuli $\forall x(\neg P(x, y) \wedge \exists y\, Q(x, y)) \wedge R(x)$, medtem ko izjavna formula $\exists y\, Q(x, y) \wedge R(x)$ v njej ne nastopa.

Formula $\forall x(\neg P(x, y) \wedge \exists y\, Q(x, y)) \wedge R(x)$ ima globino enako 4, njena dolžina pa je enaka 8. To je ravno za 1 več kot je število povezav konstrukcijskega drevesa.

Doseg kvantifikatorja in vezava spremenljivk

Natančen pomen kvantifikatorjev bomo spoznali v naslednjem razdelku, v tem razdelku pa si bomo ogledali delovanje kvantifikatorja na spremenljivkah izjavne formule. Dogovorimo se še za terminološko oznako. Spremenljivki x, ki stoji neposredno za kvantifikatorjem v zapisu $\forall x$ ali $\exists x$, bomo rekli tudi kvantifikatorju *lastna spremenljivka*.

Oglejmo si za začetek aritmetično formulo $x+5$. Njena (numerična) vrednost bo določena šele z izbiro vrednosti spremenljivke x. Pravimo tudi, da je vrednost formule *odvisna* od spremenljivke x. Izraz

$$\lim_{x \to 2}(x + 5)$$

od spremenljivke x ni odvisen (tako enostavne limite pri matematiki ne boste pogosto srečali, njena vrednost je seveda enaka 7), čeprav spremenljivka x nastopa tudi v tem izrazu. Uporaba limite (ko gre x proti 2) *deluje* na spremenljivko x v izrazu $x + 5$.

Tudi v predikatnem računu bomo opisali delovanje kvantifikatorja na spremenljivkah. Posamezen *vstop* v formuli W je bodisi *prost* ali *vezan*. Z induktivno definicijo gre takole:

(V1) vsi vstopi spremenljivk v atomarnih formulah so prosti,

(V2) v formuli $\forall x(W)$ univerzalni kvantifikator $\forall$ veže vse proste vstope spremenljivke x v formuli W,

(V3) v formuli $\exists x(W)$ eksistenčni kvantifikator $\exists$ veže vse proste vstope spremenljivke x v formuli W,

(V4) posamezen vstop spremenljivke je lahko vezan z največ enim kvantifikatorjem.

Neposredno za kvantifikatorji zapisane lastne spremenljivke[2] bomo obravnavali kot vezane.

Prednost kvantifikatorja je v skladu z dogovorom enaka kot prednost negacije. Tako v formuli

$$\forall x\, P(x) \wedge Q(x, y) \tag{2.4}$$

najprej *izvedemo* kvantifikator, šele nato uporabimo konjunkcijo. *Doseg kvantifikatorja* je torej najkrajša izjavna formula, ki jo preberemo neposredno za lastno spremenljivko

[2] Namesto o lastnih spremenljivkah bi lahko govorili tudi o dodatni členitvi kvantifikatorjev. Tako bi, denimo, obstajalo več podvrst univerzalnega kvantifikatorja, glede na ime spremenljivke, ki jo vsebujejo.

posameznega kvantifikatorja. V primeru formule (2.4) je doseg (edinega) kvantifikatorja enak formuli $P(x)$, kar prikažemo tudi takole

$$\underset{\llcorner\lrcorner\llcorner\lrcorner}{\forall \boldsymbol{x}\ P(\boldsymbol{x})} \wedge Q(x,y). \tag{2.5}$$

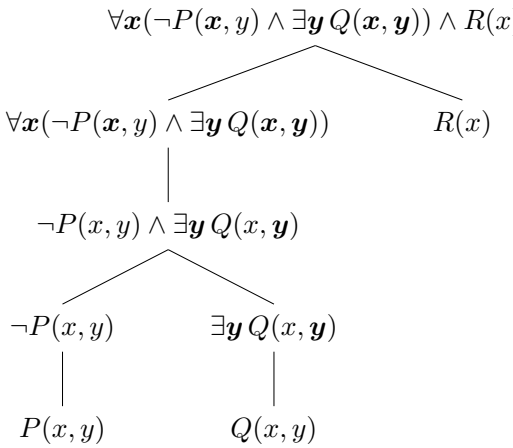
Pri tem smo vezane vstope spremenljivk v formuli označili s polkrepko pisavo.

Komentirajmo na kratko definicijo prostih in vezanih vstopov spremenljivke v formuli. Status vstopa spremenljivke najlaže spremljamo vzdolž konstrukcijskega drevesa formule. Intuitivno bodo spremenljivke v začetku gradnje izjavne formule proste (V1). Z uporabo kvantifikatorjev lahko prost vstop spremenljivke pretvorimo v vezanega. V opisih (V2) in (V3) smo z uporabo oklepajev (W) natanko opisali doseg uporabljenega kvantifikatorja. Ko je vstop spremenljivke enkrat vezan s kvantifikatorjem, naslednji kvantifikatorji takšne spremenljivke, natančneje njenega vstopa, ne vežejo ponovno (V4), četudi morda pripada njihovemu dosegu.

Oglejmo si doseg kvantifikatorjev v že opazovani formuli

$$\underset{\llcorner\lrcorner\llcorner\!\!\llcorner\!\!\lrcorner\!\!\lrcorner}{\forall x\,(\neg P(x,y) \wedge \underset{\llcorner\lrcorner\llcorner\!\!\lrcorner}{\exists y\, Q(x,y)})} \wedge R(x) \tag{2.6}$$

in opazujmo induktivno vezavo posameznih vstopov spremenljivk vzdolž njenega konstrukcijskega drevesa, glej sliko 2.2. Kakšen je status posameznih vstopov spremenljivk



Slika 2.2: Vezava spremenljivk v formuli $\forall x(\neg P(x,y) \wedge \exists y\, Q(x,y)) \wedge R(x)$.

v formuli (2.6)? Spremenljivka x v formuli nastopa štirikrat. Najprej kot lastna (in zato vezana) spremenljivka edinega univerzalnega kvantifikatorja, sledita dva vezana vstopa spremenljivke x, na koncu pa še en prost vstop le-te. Spremenljivka y se vzdolž formule najprej pojavi kot prosta spremenljivka, nato kot lastna spremenljivka eksistenčnega kvantifikatorja in slednjič kot vezana spremenljivka.

Še dve kratki, a povezani definiciji. Izjavno formulo brez prostih spremenljivk imenujemo tudi *izjavna shema*, včasih tudi *zaprta formula*. Izjavno formulo, v kateri ne uporabimo nobenega kvantifikatorja, imenujemo *odprta izjavna formula*.

2.3 Interpretacija izjavne formule

V tem razdelku se bomo ukvarjali z vprašanjem, kako izjavno formulo spremeniti v izjavo. Formula $\forall x\, P(x)$ ne vsebuje prostih spremenljivk, toda logična vrednost izraza je seveda odvisna od pomena predikata P.

Interpretacija $\mathcal{I}$ izjavne formule W je sestavljena z izbiro

(I1) *področja pogovora* interpretacije, ki je neprazna družina $\mathcal{D}$, iz katere jemljemo posamezne konstante,

(I2) *pomena predikatov* — vsakemu k-mestnemu predikatu P, ki ga izjavna formula W uporablja, ustreza logična preslikava, definirana na družini k-teric elementov iz $\mathcal{D}$,

(I3) *pomena konstant* — vsaki konstanti ustreza natančno določen element področja pogovora, in

(I4) *izbire prostih spremenljivk* — vsaki prosti spremenljivki izberemo vrednost iz $\mathcal{D}$.

Za izračun logične vrednosti izjavne formule W glede na izbrano interpretacijo moramo definirati še pomen kvantifikatorjev. Naj bo W izjavna formula in t spremenljivka ali konstanta, term z eno besedo. Oznaka $W(x/t)$ označuje formulo, ki jo dobimo, če vse proste vstope spremenljivke x nadomestimo s t. Pravimo, da smo izjavno formulo $W(x/t)$ dobili z *zamenjavo* oziroma *substitucijo* spremenljivke x s termom t.

Za formulo

$$W \quad = \quad \forall x\, P(x) \wedge Q(x, y)$$

si oglejmo nekaj zamenjav spremenljivk.

$$\begin{aligned} W(x/a) &= \forall x\, P(x) \wedge Q(a, y) \\ W(x/y) &= \forall x\, P(x) \wedge Q(y, y) \\ W(x/z) &= \forall x\, P(x) \wedge Q(z, y) \\ W(y/b) &= \forall x\, P(x) \wedge Q(x, b) \end{aligned}$$

Opazimo, da zamenjava (x/t) ne vpliva na vezane vstope spremenljivke x, ravno tako ne spremeni vstopov spremenljivk z drugimi imeni.

Zamenjava spremenljivke x s termom t je načeloma možna v vsakem primeru. Natančneje pa pravimo, da je spremenljivka x *zamenljiva* s termom t, če noben prost vstop

spremenljivke x ne leži v dosegu kvantifikatorja z lastno spremenljivko t (tj. če je t konstanta, potem je x gotovo zamenljiva s t, če pa je t spremenljivka, potem x ne sme nastopati prosto v dosegu kvantifikatorja $\forall t$ ali $\exists t$). V izjavni formuli

$$W \quad = \quad \forall x\, P(x) \wedge \exists y\, Q(x, y)$$

spremenljivka x *ni zamenljiva* s spremenljivko y, saj z zamenjavo (x/y) pridelamo formulo

$$W(x/y) \quad = \quad \forall x\, P(x) \wedge \exists y\, Q(y, y).$$

Pri tem spremenljivka x v formuli W nastopa prosto, a se nahaja v dosegu kvantifikatorja z lastno spremenljivko y.

Vsekakor je v formuli W spremenljivka x zamenljiva s termom t vsaj v naslednjih dveh primerih:

- t je konstanta ali
- t je nova spremenljivka, ki v W ne nastopa.

Končno moremo definirati *pomen kvantifikatorjev* v izjavni formuli. Začnimo z univerzalnim kvantifikatorjem. Formula

$$\forall x\, (W) \tag{2.7}$$

je resnična v interpretaciji $\mathcal{I}$ s področjem pogovora $\mathcal{D}$, če je za vsak $d \in \mathcal{D}$ v tej isti interpretaciji resnična tudi formula

$$(W)(x/d).$$

Sicer je formula (2.2) neresnična.

Analogno definicijo uporabimo na eksistenčnem kvantifikatorju. Formula

$$\exists x\, (W) \tag{2.8}$$

je resnična v interpretaciji $\mathcal{I}$ s področjem pogovora $\mathcal{D}$, če v področju pogovora obstaja $d \in \mathcal{D}$, za katerega je v tej isti interpretaciji resnična tudi formula

$$(W)(x/d).$$

Sicer je formula (2.3) neresnična.

Še termin za konec razdelka. Interpretacijo $\mathcal{I}$, v kateri je izjavna formula W resnična, imenujemo tudi *model* formule W.

2.4 Enakovrednost izjavnih formul

V izjavnem računu smo spoznali *tavtologijo*, *protislovje* in pare *enakovrednih* izjavnih izrazov. Tavtologija je izjavni izraz, ki je *vedno* resničen, pri čemer besedica vedno pomeni "pri vseh naborih logičnih vrednosti v izrazu nastopajočih spremenljivk".

Ustreznice v predikatnem računu definiramo takole. Izjavna formula W je *splošno veljavna*, če ima W v vsaki interpretaciji logično vrednost 1. Izjavna formula V je *neizpolnljiva*, če je V v vsaki interpretaciji napačna.

Analogno pravimo, da sta formuli U in U' *enakovredni*, pišemo tudi $U \sim U'$, če imata U in U' v vsaki interpretaciji isto logično vrednost.

V veljavi je analog izreka 1.1 iz izjavnega računa. Navedimo ga brez dokaza.

Izrek 2.1 *Izjavni formuli U in U' sta enakovredni natanko tedaj, ko je formula $U \Leftrightarrow U'$ splošno veljavna.*

V izjavnem računu je testiranje enakovrednosti izjavnih izrazov rutinski, če že zamuden, postopek. V najslabšem primeru primerjamo pravilnostni tabeli. Sicer sta lahko dolgi, a vseeno omejene dolžine. V predikatnem računu si preverjanja vseh možnih interpretacij ne moremo privoščiti, interpretacij je namreč neskončno mnogo (že samo za področja pogovora si lahko izberemo različne podmnožice množice naravnih števil).

Edina praktična možnost preverjanja enakovrednosti formul je izpeljava ene od formul začenši iz druge formule z uporabo elementarnih enakovrednosti. Le-te so bodisi uporabe zakonov izjavnega računa na nivoju izjavnih formul ali pa predikatnemu računu lastne enakovrednosti, ki jih bomo imenovali tudi zakoni predikatnega računa.

Prepis iz izjavnega računa

Izjavna izraza

$$p \Rightarrow (q \Rightarrow p) \quad \text{in} \quad p \wedge \neg p$$

sta po vrsti tavtologija in protislovje. Denimo, da namesto izjavne spremenljivke p zapišemo formulo $\forall x\, P(x)$ in namesto spremenljivke q formulo $\exists y\, Q(x, y)$. Pridelamo formuli

$$\forall x\, P(x) \Rightarrow (\exists y\, Q(x, y) \Rightarrow \forall x\, P(x)) \quad \text{in} \quad \forall x\, P(x) \wedge \neg \forall x\, P(x).$$

Prva od obeh formul je splošno veljavna, druga je neizpolnljiva.

Podoben trik lahko izvedemo s parom enakovrednih izjavnih izrazov, denimo,

$$p \Rightarrow q \quad \text{in} \quad \neg p \vee q.$$

Z uporabo zgoraj predstavljenega prepisa pridelamo par enakovrednih izjavnih formul

$$\forall x\, P(x) \Rightarrow \exists y\, Q(x, y) \quad \text{in} \quad \neg \forall x\, P(x) \vee \exists y\, Q(x, y).$$

Vsak par enkovrednih izjavnih formul lahko s prepisom — nadomeščanjem izjavnih spremenljivk z izjavnimi formulami — spremenimo v par enakovrednih izjavnih formul, tavtologije oziroma protislovja pa prepišemo v splošno veljavne oziroma neizpolnljive izjavne formule.

Zakoni predikatnega računa

V tem in naslednjem razdelku bomo iskali pare enakovrednih izjavnih formul, ki so lastni predikatnemu računu. Posebej enostavne pare tipov enakovrednih izjavnih formul (ki niso prepisani iz enakovrednosti izjavnega računa) bomo imenovali *zakone predikatnega računa*.

V tem razdelku navedemo tri družine zakonov predikatnega računa, ki veljajo brezpogojno.

(1) **de Morganova zakona:**

$$\neg\forall x\,(W) \quad\sim\quad \exists x\neg(W) \quad \text{in} \quad \neg\exists x\,(W) \quad\sim\quad \forall x\neg(W)$$

(2) **zamenjava istovrstnih kvantifikatorjev:**

$$\forall x\,\forall y\,(W) \quad\sim\quad \forall y\,\forall x\,(W) \quad \text{in} \quad \exists x\,\exists y\,(W) \quad\sim\quad \exists y\,\exists x\,(W)$$

(3) **distributivnost:**

$$\forall x(V \wedge W) \quad\sim\quad \forall x\,(V) \wedge \forall x\,(W) \quad \text{in} \quad \exists x(V \vee W) \quad\sim\quad \exists x\,(V) \vee \exists x\,(W)$$

Ilustrirajmo eno verzijo de Morganovega zakona vsaj v eni smeri. Izberimo model $\mathcal{I}$ za formulo $\neg\forall x\,(W)$. V tej isti interpretaciji je izjava $\forall x\,(W)$ lažna, kar pomeni, da $W(x/d)$ ni resnična za vsak d iz področja pogovora interpretacije. To pomeni, da v področju pogovora obstaja konkreten element d, za katerega izjava $W(x/d)$ ni resnična, oziroma je resnična izjava $\neg W(x/d)$. Po definiciji eksistenčnega kvantifikatorja je torej resnična izjava $\exists x\neg(W)$. Torej je $\mathcal{I}$ tudi model za $\exists x\neg(W)$.

Zamenjavo istovrstnih kvantifikatorjev (2) utemeljimo takole. Pri izbrani interpretaciji s področjem pogovora $\mathcal{D}$ je za vsak par elementov področja pogovora d, d' izraz $(W)(x/d)(y/d')$ enak izrazu $(W)(y/d')(x/d)$. Torej imata

$$(W)(x/d)(y/d') \quad \text{in} \quad (W)(y/d')(x/d)$$

isto logično vrednost.

V splošnem pa ne smemo zamenjati vrstnega reda različnih kvantifikatorjev. Izjavni formuli

$$\forall x\exists y\,P(x,y) \quad \text{in} \quad \exists y\forall x\,P(x,y) \tag{2.9}$$

nista enakovredni. Interpretacija v področju pogovora naravnih števil, kjer predikat $P(x, y)$ beremo kot *"x je manjši ali enak y"*, je model za natančno eno od obeh formul. Bralec naj sam preveri, katera izmed formul (2.9) je v tej interpretaciji resnična.

Zakon (3) trdi, da univerzalni kvantifikator distribuira preko konjunkcije in eksistenčni preko disjunkcije. Model za izjavno formulo $\forall x\,(V \wedge W)$ je namreč tudi model za formuli $\forall x\,(V)$ in $\forall x\,(W)$. Za drugo smer premisleka izberimo model $\mathcal{I}$ za formulo $\forall x\,(V) \wedge \forall x\,(W)$. V tej interpretaciji sta resnični tako $\forall x\,(V)$ kot $\forall x\,(W)$, kar pomeni, da je za vsak element področja pogovora d resnična izjava $(V)(x/d)$ pa tudi izjava $(W)(x/d)$. Ker je $((V)(x/d)) \wedge ((W)(x/d)) = (V \wedge W)(x/d)$, je $\mathcal{I}$ tudi model za izjavno formulo $\forall x\,(V \wedge W)$.

Po drugi strani univerzalni kvantifikator ne distribuira preko disjunkcije, ravno tako kot eksistenčni ne distribuira preko konjunkcije. Za področje pogovora izberimo množico naravnih števil, predikata $P(x)$ in $Q(x)$ naj po vrsti pomenita *"x je sodo število"* in *"x je liho število"*. Obstajajo tako soda kot liha naravna stevila, ne obstaja pa število, ki bi bilo liho in sodo hkrati. V tej interpretaciji je torej izjavna formula

$$\exists x\, P(x) \wedge \exists x\, Q(x)$$

resnična, medtem ko je izjavna formula

$$\exists x(P(x) \wedge Q(x))$$

lažna.

Zakoni predikatnega računa z omejitvami

Nekatere enakovrednosti v predikatnem računu so pogojene s strukturo posameznih formul. Pri vsakem od naslednjih zakonov bomo navedli pogoje, pod katerimi je posamezen zakon veljaven. Zakoni bodo, kot v predhodnem razdelku, vsaj deloma opremljeni s kratkimi idejami utemeljitev. Ravno tako bomo na kratko komentirali, zakaj so omejitve potrebne.

Obnašali se bomo ekološko. Pri ilustraciji potrebnosti zapisanih omejitev bomo v dveh primerih reciklirali isto interpretacijo. Področje pogovora naj bo množica naravnih števil, predikat $P(x)$ naj pomeni *"x je praštevilo"*, za vrednost morebitne proste spremenljivke x ali y pa izberimo število 5.

(4) **odvečni kvantifikator:** če spremenljivka x ne nastopa prosto v formuli W, potem veljata enakovrednosti

$$\forall x\,(W) \;\sim\; W \quad \text{in} \quad \exists x\,(W) \;\sim\; W$$

Ideja utemeljitve gre takole. Naj bo $\mathcal{I}$ model za formulo $\forall x\,(W)$. To pomeni, da je za vsak element področja pogovora d izjavna formula $(W)(x/d)$ resnična. Ker W ne vsebuje prostih vstopov spremenljivke x, je $W = (W)(x/d)$, zato je $\mathcal{I}$ tudi model za W.

Formuli $\forall x\, P(x)$ in $P(x)$ nista enakovredni, saj imata v zgoraj opisani interpretaciji različni logični vrednosti. Formuli $\forall x\, P(y)$ in $P(y)$ pa sta enakovredni, saj x ne nastopa prosto v $P(y)$.

(5) **preimenovanje spremenljivk:** če je v formuli W spremenljivka x zamenljiva s spremenljivko y, potem veljata enakovrednosti

$$\forall x\,(W) \;\sim\; \forall y(W(x/y)) \quad \text{in} \quad \exists x\,(W) \;\sim\; \exists y(W(x/y))$$

Naj bo $\mathcal{I}$ model za formulo $\forall x\,(W)$. To pomeni, da je $(W)(x/d)$ resnična za vsak element področja pogovora d. Toda formula $W(x/d)$ je identična formuli $W(x/y)(y/d)$, saj vsakemu prostemu vstopu spremenljivke x v formuli W ustreza vstop proste spremenljivke y v formuli $W(x/y)$. To pomeni, da je $\mathcal{I}$ tudi model za $\forall y(W(x/y))$.

V formuli $\forall y\, Q(x,y)$ spremenljivka x *ni* zamenljiva s spremenljivko y, in tudi formuli $\forall x \forall y\, Q(x,y)$ in $\forall y \forall y\, Q(y,y)$ nista enakovredni. Če izberemo področje pogovora naravnih števil in za predikat $Q(x,y)$ izberemo pomen *"x je enako y"*, je prva izmed formul neresnična, medtem ko je druga, sicer enakovredna formuli $\forall y\, Q(y,y)$, resnična.

(6) **kvantifikator in konjunkcija/disjunkcija:** če spremenljivka x ne nastopa prosto v formuli C, potem veljajo enakovrednosti

$$\begin{array}{lcl@{\qquad\qquad}lcl} C \wedge \forall x\,(W) & \sim & \forall x(C \wedge W) & C \vee \forall x\,(W) & \sim & \forall x(C \vee W) \\ C \wedge \exists x\,(W) & \sim & \exists x(C \wedge W) & C \vee \exists x\,(W) & \sim & \exists x(C \vee W) \end{array}$$

Dokazujemo samo enostransko in samo prvega od štirih primerov. Naj bo $\mathcal{I}$ model za formulo $\forall x(C \wedge W)$. To pomeni, da je za vsak d področja pogovora $(C \wedge W)(x/d)$ resnična izjava. Toda $(C \wedge W)(x/d) = C \wedge (W(x/d))$, saj C ne vsebuje prostih vstopov spremenljivke x. To pomeni, da je $\mathcal{I}$ tako model za C kot tudi model za $\forall x\,(W)$.

Zakaj je pogoj o nenastopanju prostega vstopa spremenljivke res potreben? Spet reciklaža interpretacije. Formula $\forall x(P(x) \vee P(x))$ je v tej interpretaciji lažna, medtem ko je formula $P(x) \vee \forall x\, P(x)$ resnična.

Zakone predikatnega računa z omejitvami (6) smemo prebrati kot možnost vpeljave kvantifikatorja (kateregakoli) v notranjost konjunkcije oziroma disjunkcije, če kateri izmed členov kvantifikatorja *ne potrebuje*. Ravno tako lahko omenjene zakone preberemo v *nasprotno smer*, kako lahko doseg kvantifikatorja prestavimo s posameznega člena konjunkcije oz. disjunkcije na celoten izraz, če je le preostali člen *imun* nanj.

Ali obstajajo ustreznice zakonov, navedenih v točki (6), za primer implikacije? Da in ne, je pravilni odgovor. Tudi za implikacijo smemo zapisati ustrezne zakone za prenos kvantifikatorja k enemu samemu členu implikacije. Je pa res, da tega ne bomo storili. Razlog je v nekomutativnosti implikacije, člena v implikaciji imata namreč različni vlogi. To

pomeni, da bi za vsakega od kvantifikatorjev potrebovali kar dve pravili. Po drugi strani pravil ne potrebujemo, saj lahko implikacijo enakovredno prepišemo kot diskunkcijo.

Vseeno predpostavimo, da spremenljivka x v formuli C ne nastopa prosto, in zapišimo:

$$\begin{aligned} \forall x\,(W) \Rightarrow C \quad &\sim \quad \neg\forall x\,(W) \vee C \\ &\sim \quad \exists x \neg(W) \vee C \\ &\sim \quad \exists x(\neg W \vee C) \\ &\sim \quad \exists x(W \Rightarrow C) \end{aligned}$$

Pri prenosu kvantifikatorja k posameznemu členu implikacije se lahko vrsta kvantifikatorja spremeni.

2.5 Preneksna normalna oblika

Preneksna normalna oblika formul W je izjavna formula W_{PNO}, za katero velja

(PNO1) $W_{PNO} \sim W$ in

(PNO2) v formuli W_{PNO} se vsi kvantifikatorji nahajajo neposredno na začetku formule.

Formula

$$\forall x\, P(x) \wedge \exists y\, Q(y)$$

je enakovredna formuli

$$\forall x \exists y (P(x) \wedge Q(y)),$$

pri čemer je slednja zapisana v preneksni normalni obliki.

Preneksna normalna oblika formule je pomembna za dokazovanje enakovrednosti formul z izpeljavo. Če znamo obe (domnevno enakovredni) formuli zapisati v preneksni normalni obliki, bi lahko enakovrednost dokazali s primerjanjem vrstnega reda kvantifikatorjev v začetku formule, enakovrednost odprtih formul v preostanku pa bi poskusili pokazati s prevedbo samo z uporabo zakonov izjavnega računa.

K sreči lahko vsako izjavno formulo zapišemo v preneksni normalni obliki.

Izrek 2.2 *Vsako izjavno formulo W lahko enakovredno prepišemo v* preneksni normalni obliki *s kvantifikatorji na neposrednem začetku.*

Formalnega dokaza izreka 2.2 ne bomo izdelali. Idejo si bomo ogledali na konkretnem zgledu, izjavni formuli

$$\forall z(P(z) \Rightarrow \neg\exists y(Q(z,y) \wedge \forall z R(y,z))). \tag{2.10}$$

De Morganov zakon predikatnega računa (1) omogoča prenos negacije preko kvantifikatorja. Distributivnost (3) in zakoni z omejitvami (6) omogočajo prenos kvantifikatorjev od posameznega člena disjunkcije ozirom konjunkcije proti začetku izjavne formule. Prvi korak je torej

(PX1) **Zamenjava izjavnih veznikov:** izjavne veznike nadomestimo s konjunkcijami, disjunkcijami in negacijami.

Ker je $\{\neg, \wedge, \vee\}$ poln nabor izjavnih veznikov, nam korak (PX1) uspe. V formuli (2.10) nastopa implikacija, enakovredna formula se glasi

$$\forall z(\neg P(z) \vee \neg \exists y(Q(z,y) \wedge \forall z R(y,z))). \tag{2.11}$$

Naslednji korak zapisa formule v preneksni normalni obliki poskrbi za preimenovanje spremenljivk.

(PX2) **Preimenovanje spremenljivk:** nobena spremenljivka (natančneje njeno ime) ne nastopa hkrati vezano in prosto, ravno tako ne sme biti vezana z več kot enim kvantifikatorjem.

V izjavni formuli (2.11) spremenljivka z imenom z nastopa v dvojni vlogi, dva kvantifikatorja vežeta njene vstope. V skladu z zakonom predikatnega računa (6) lahko formulo $\forall z(W)$ enakovredno prepišemo kot $\forall x(W(z/x))$, če je spremenljivka z zamenljiva s spremenljivko x, glej zakon predikatnega računa (5). Ker spremenljivka x v (2.11) sploh ne nastopa, lahko formulo enakovredno prepišemo v

$$\forall x(\neg P(x) \vee \neg \exists y(Q(x,y) \wedge \forall z R(y,z))). \tag{2.12}$$

V formuli (2.11) je spremenljivka z vezana z dvema kvantifikatorjema, s preimenovanjem spremenljivk smo ta pojav v formuli (2.12) odpravili.

Sledi

(PX3) **Prenos kvantifikatorjev proti začetku:** z uporabo zakonov predikatnega računa (1), (3) in (6) uspemo kvantifikatorje prestaviti proti začetku izjavne formule.

Bomo res lahko uporabili zakon z omejitvami (6)? Če smo spremenljivke preimenovali v skladu z (PX2), velja naslednja lastnost. Če kvantifikator veže posamezen člen konjunkcije oziroma disjunkcije, potem njegova lastna spremenljivka v preostalih členih konjunkije oziroma disjunkcije ne nastopa. Takšen kvantifikator lahko po (6) prestavimo na začetek celotne konjunkcije oziroma disjunkcije. Računajmo

$$\begin{aligned}
\forall x(\neg P(x) \vee \neg \exists y(Q(x,y) \wedge \forall z\, R(y,z))) \quad &\sim \quad \forall x(\neg P(x) \vee \forall y \neg (Q(x,y) \wedge \forall z\, R(y,z))) \\
&\sim \quad \forall x(\neg P(x) \vee \forall y(\neg Q(x,y) \vee \neg \forall z\, R(y,z))) \\
&\sim \quad \forall x \forall y(\neg P(x) \vee \neg Q(x,y) \vee \neg \forall z\, R(y,z)) \\
&\sim \quad \forall x \forall y(\neg P(x) \vee \neg Q(x,y) \vee \exists z \neg R(y,z)) \\
&\sim \quad \forall x \forall y \exists z(\neg P(x) \vee \neg Q(x,y) \vee \neg R(y,z))
\end{aligned}$$

Rezultat je izjavna formula v preneksni normalni obliki, kar zaključuje ilustracijo izreka 2.2.

V začetku razdelka smo namignili, da lahko enakovrednost izjavnih formul preverjamo s prevedbo formul v preneksno normalno obliko, potem pa po eni strani preverimo ujemanje začetnega dela formule s kvantifikatorji in tudi enakovrednost odprtih formul brez kvantifikatorjev. Seveda je potrebno paziti na imena uporabljenih spremenljivk, ravno tako pa na vrstni red kvantifikatorjev. V splošnem, glej zvezo (2.9), raznovrstnih kvantifikatorjev v formuli ne smemo zamenjati. Včasih, v odvisnosti od zgradbe odprtega dela formule, pač.

Velja namreč enakovrednost

$$\forall x \exists y (P(x) \wedge Q(y)) \quad \sim \quad \exists y \forall x (P(x) \wedge Q(y)). \tag{2.13}$$

Računajmo:

$$\begin{aligned} \forall x \exists y (P(x) \wedge Q(y)) \quad &\sim \quad \forall x (P(x) \wedge \exists y\, Q(y)) \\ &\sim \quad \forall x\, P(x) \wedge \exists y\, Q(y) \\ &\sim \quad \exists y (\forall x\, P(x) \wedge Q(y)) \\ &\sim \quad \exists y \forall x (P(x) \wedge Q(y)) \end{aligned}$$

Pri tem smo kvantifikatorja najprej poslali v globino izjavne formule, nato pa znova proti začetku. Ker kvantifikatorja $\forall x$ in $\exists y$ delujeta na različnih členih konjunkcije v izrazu $\forall x\, P(x) \wedge \exists y\, Q(y)$, sta obe možnosti, katerega od obeh kvantifikatorjev prednostno prestavimo na začetek, enako dobri.

2.6 Sklepanje v predikatnem računu

Začnimo z zgledom. Pokažimo, da je izjavna formula

$$\neg \exists y \forall x (P(x, y) \Leftrightarrow \neg P(x, x)) \tag{2.14}$$

splošno veljavna.

Pokazati je potrebno, da je formula (2.14) resnična v vsaki možni interpretaciji. Uporabljali bomo dokaz s protislovjem. Privzeli bomo, da obstaja interpretacija $\mathcal{I}$ s področjem pogovora $\mathcal{D}$, v kateri formula (2.14) ni resnična.

Interpretacija $\mathcal{I}$ je torej model za formulo

$$\exists y \forall x (P(x, y) \Leftrightarrow \neg P(x, x)),$$

ki je negacija originalne formule. Pomen eksistenčnega kvantifikatorja trdi, da v področju pogovora obstaja element $d \in \mathcal{D}$, za katerega je izjava

$$\forall x (P(x, d) \Leftrightarrow \neg P(x, x))$$

resnična. To pomeni, da je za vsak $d' \in \mathcal{D}$ resnična izjava

$$P(d', d) \Leftrightarrow \neg P(d', d').$$

To pa ni možno, saj z izbiro $d' = d$ pridelamo protislovje

$$P(d, d) \Leftrightarrow \neg P(d, d)$$

Torej ne obstaja interpretacija, v kateri je formula (2.14) lažna. Oziroma, formula (2.14) je splošno veljavna.

Formula (2.14) sicer ni zapisana v preneksni normalni obliki. Njena negacija, ki jo uporabimo kot predpostavko sklepa s protislovjem, pač. Kvantifikatorje, zapisane v začetku formule, smo znali odpraviti in pridelati odprto formulo. V tej utemeljitvi smo končali takoj, ko smo pridelali protislovje. V splošnem pa bomo želeli kvantifikatorje uvesti nazaj.

Poleg pravil sklepanja iz izjavnega računa bomo v predikatnem računu uporabljali tudi pravila, ki po eni strani zmorejo odpraviti kvantifikatorje, po drugi strani pa jih bodo znali tudi uvesti nazaj v formulo.

Začnimo z izrekom, ki pravilen sklep v predikatnem računu enači s splošno veljavnostjo formule. Gre za analog izreka 1.12 iz izjavnega računa.

Izrek 2.3 *$W_1, W_2, \ldots, W_n \models W$ natanko tedaj, ko je izjavna formula $(W_1 \wedge W_2 \wedge \ldots \wedge W_n) \Rightarrow W$ tavtologija.*

Dokaz. Izjavna formula $(W_1 \wedge W_2 \wedge \ldots \wedge W_n) \Rightarrow W$ je splošno veljavna natanko tedaj, ko je resnična v vsaki interpretaciji $\mathcal{I}$. To pomeni, da je v vsaki interpretaciji, ki je model za vse predpostavke $W_1, W_2, \ldots, W_n$, resničen tudi zaključek W. □

Pravila sklepanja

Pravila sklepanja izjavnega računa so v veljavi tudi v predikatnem računu. Premisliti je potrebno, da je vsaka interpretacija, ki je model za vse predpostavke (v kateri so vse predpostavke resnične), tudi model zaključka.

Vzemimo pod drobnogled pravilo sklepanja *modus ponens*

$$A, A \Rightarrow B \models B$$

in izberimo interpretacijo $\mathcal{I}$, ki je model obeh predpostavk A in $A \Rightarrow B$. Po definiciji implikacije je v tem primeru $\mathcal{I}$ tudi model za formulo B. Na podoben način ukanimo preostalih šest pravil sklepanja iz izjavnega računa, predstavljenih v tabeli 1.6.

Predikatnemu računu lastna pravila sklepanja bodo vključevala kvantifikatorje. Za vsakega od kvantifikatorjev, univerzalnega in eksistenčnega, bomo spoznali par pravil. Pravili, ki kvantifikator odpravita, bomo imenovali specifikaciji. Pravili, ki bosta kvantifikator uvedli, imenujemo generalizaciji.

Univerzalna specifikacija US

$$\forall x(W) \models W(x/t) \qquad \text{če je } x \text{ zamenljiva s } t. \tag{2.15}$$

Pogoj, da je spremenljivka x zamenljiva s t, je gotovo izpolnjen, če je t konstanta ali $t = x$.

Intuitivno univerzalno specifikacijo razumemo takole. Če je veljavna formula $\forall x(W)$, potem je W veljavna za vsak element področja pogovora, ki ga vstavimo namesto spremenljivke x. Pri zamenjavi spremenljivke x s spremenljivko y pa je potrebno paziti, da vsakemu prostemu vstopu spremenljivke x v formuli W ustreza prost vstop spremenljivke y v formuli $W(x/y)$. To zahtevo kršimo v primeru, ko prost vstop x v formuli W leži v dosegu kvantifikatorja z lastno spremenljivko y. Pri zamenjavi (x/y) bi bil dobljeni vstop spremenljivke y vezan.

Dualno pravilo sklepanja je eksistenčna generalizacija.

Eksistenčna generalizacija EG

$$W(x/t) \models \exists x\, W \qquad \text{če je } x \text{ zamenljiva s } t. \tag{2.16}$$

Implicitno privzamemo, da je x prosta spremenljivka v formuli W. V nasprotnem primeru je formula $W(x/t)$ enaka W, in je tudi enakovredna formuli $\exists x\, W$.

Pogoj, da je spremenljivka x zamenljiva s t, je gotovo izpolnjen, če je t konstanta ali $t = x$.

Namesto dokaza pravilnosti univerzalne specifikacije in eksistenčne generalizacije raje pokažimo, da gre res za dualno zapisani pravili. Privzemimo, da je v formuli W spremenljivka x res zamenljiva s t. Univerzalna specifikacija, kot je zapisana v (2.15), je pravilen sklep natanko tedaj, ko je formula

$$\forall x(W) \Rightarrow (W(x/t))$$

splošno veljavna. Računajmo:

$$\begin{aligned} \forall x(W) \Rightarrow (W(x/t)) \quad &\sim \quad \neg(W(x/t)) \Rightarrow \neg\forall x(W) \\ &\sim \quad \neg(W(x/t)) \Rightarrow \exists x \neg(W) \\ &\sim \quad U(x/t) \Rightarrow \exists x\, U \end{aligned}$$

Pri tem smo v zadnji enakovrednosti negacijo formule W označili kot U in pridelali implikacijo, ki opisuje eksistenčno generalizacijo.

Nekoliko več težav je z preostalima dvema praviloma sklepanja, odpravo eksistenčnega kvantifikatorja in uvedbo univerzalnega kvantifikatorja.

Eksistenčna specifikacija ES

$$\exists x(W) \models W(x/c) \qquad \begin{array}{l}\text{če je } c \text{ konstanta, ki}\\ \bullet \text{ je } \textit{nova} \text{ in}\\ \bullet \text{ ne nastopa v zaključku sklepa.}\end{array} \tag{2.17}$$

Izbira konkretnega elementa področja pogovora, za katerega je resničen zaključek $W(x/c)$, je odvisna od vseh izbir za proste spremenljivke, ki nastopajo v formuli $\exists x(W)$. Če so $y_1, \ldots, y_k$ natanko vse spremenljivke, ki prosto nastopajo v predpostavki $\exists x(W)$, potem pravimo, da je c *odvisna* od spremenljivk $y_1, \ldots, y_k$ oziroma tudi, da je $c = c(y_1, \ldots, y_k)$ *Skolemova funkcija* spremenljivk $y_1, \ldots, y_k$.

Pogoj eksistenčne specifikacije je gotovo potreben. Pravilnost formule $\exists x(W)$ pomeni, da v področju pogovora obstaja element c, za katerega je resnična formula $W(x/c)$. Nobenega razloga ni, da bi c bil enak kateremu od predhodno izpostavljenih elementov, bodisi v prehodnem teku dokaza bodisi v zaključku. Tako iz predpostavke $\exists x P(x, c)$ ne moremo sklepati na, denimo, $P(c, c)$. Lahko pa, v primeru, da je d *nova* konstanta, sklepamo na $P(d, c)$.

Univerzalna generalizacija UG

$$W \models \forall x(W) \qquad \begin{array}{l}\bullet \text{ če } x \text{ ne nastopa prosto v nobeni od predpostavk in}\\ \bullet \text{ v dokazu } W \text{ spremenljivka } x \text{ ni argument Skolemove funkcije.}\end{array} \tag{2.18}$$

Prostim spremenljivkam v predpostavkah se lahko izognemo tako, da jih predhodno nadomestimo s konstantami. Interpretacija družine formul namreč zahteva, da za proste spremenljivke izberemo konkretne elemente področja pogovora. Spremenljivka x je argument Skolemove funkcije $c(x)$, če x nastopa prosto v formuli $\exists y(U)$, ki smo jo uporabili v pravilu eksistenčne specifikacije, in sklepali na formulo $U(y/c)$.

Zgleda pravilnih sklepov

Zamenjava raznovrstnih kvantifikatorjev ne ohranja formul po enakovrednosti. Vseeno pa velja implikacija v eno smer, prenos eksistenčnega kvantifikatorja v globino formule je možen.

$$\exists y \forall x P(x, y) \models \forall x \exists y P(x, y)$$

1.	$\exists y \forall x P(x, y)$	predpostavka
2.	$\forall x P(x, c)$	ES(1)
3.	$P(x, c)$	US(2)
4.	$\exists y P(x, y)$	EG(3)
5.	$\forall x \exists y P(x, y)$	UG(4)

Vemo, da lahko univerzalni kvantifikator distribuiramo vzdolž členov konjunkcije. Univerzalni kvantifikator, ki se pojavi pri več členih disjunkcije, smemo "izpostaviti".

$$\forall x\, P(x) \vee \forall x\, Q(x) \;\models\; \forall x(P(x) \vee Q(x))$$

1.	$\forall x\, P(x) \vee \forall x\, Q(x)$	predpostavka
2.	$\forall x\, P(x) \vee \forall y\, Q(y)$	$\sim$(1)
3.	$\forall x(P(x) \vee \forall y\, Q(y))$	$\sim$(2)
4.	$\forall x \forall y(P(x) \vee Q(y))$	$\sim$(3)
5.	$\forall y(P(x) \vee Q(y))$	US(4)
6.	$P(x) \vee Q(x)$	US(5)(y/x)
7.	$\forall x(P(x) \vee Q(x))$	UG(6)

Zgledi nepravilnih sklepov

V zadnjem razdelku ilustriramo, zakaj so pogoji, opisani pri pravilih sklepanja, res potrebni. Predstavili bomo nekaj napačnih sklepov. Za vsakega bomo najprej poiskali interpretacijo, ki je model za predpostavke in ni model zaključka, nato pa zapisali neuspešen poskus dokaza pravilnosti.

Napačen sklep: $\exists y \exists x P(x, y) \;\models\; \exists x P(x, x)$

Sklep ni pravilen — v področju pogovora ljudi *"obstajata človeka, od katerih je eden oče drugemu"*, toda *"ne obstaja človek, ki je oče samemu sebi."*

1.	$\exists y \exists x P(x, y)$	predpostavka	
2.	$\exists x P(x, c)$	ES(1)	
3.	$\exists x \exists x P(x, x)$	EG(2)	NAPAKA!!
4.	$\exists x P(x, x)$	$\sim$(3)	

Napaka je skrita v 3. vrstici. Spremenljivka x ne more nadomestiti konstante c v formuli $\exists x P(x, c)$, saj se c nahaja v dosegu eksistenčnega kvantifikatorja z lastno spremenljivko x.

Napačen sklep: $\exists x P(x) \wedge \exists x Q(x) \;\models\; \exists x(P(x) \wedge Q(x))$

Področje pogovora so naravna števila, predikata $P(x)$ in $Q(x)$ naj pomenita *"x je sodo število"* oziroma *"x je liho število."* Obstajajo tako liha kot soda naravna števila, zato je predpostavka resnična, medtem ko nobeno naravno število ni hkrati liho in sodo. Zaključek ima v tej interpretaciji logično vrednost 0, zato je sklep napačen.

1.	$\exists x P(x) \wedge \exists x Q(x)$	predpostavka
2.	$\exists x P(x)$	Po(1)

3.	$\exists x Q(x)$	Po(1)	
4.	$P(c)$	ES(2)	
5.	$Q(c)$	ES(2)	NAPAKA!!
6.	$P(c) \wedge Q(c)$	Zd(4,5)	
7.	$\exists x(P(x) \wedge Q(x))$	EG(6)	

V 5. vrstici gre za napačno rabo eksistenčne specifikacije, potrebovali bi novo konstanto.

Napačen sklep: $P(x) \models \forall x\, P(x)$

Področje pogovora so naravna števila, predikat $P(x)$ naj pomeni *"x je praštevilo"*. Za vrednost proste spremenljivke x izberimo število 7. Predpostavka $P(x)$ je v tej interpretaciji resnična, saj je 7 res praštevilo. Toda zaključek je napačen, saj niso vsa naravna števila tudi praštevila.

1.	$P(x)$	predpostavka	
2.	$\forall x P(x)$	UG(1)	NAPAKA!!

V 2. vrstici uporabljeno pravilo UG zahteva, da spremenljivka x, ker uvajamo univerzalni kvantifikator z lastno spremenljivko x, ne nastopa prosto v predpostavkah.

Napačen sklep: $\forall x \exists y P(x,y) \models \exists y \forall x P(x,y)$

Pogovarjajmo se o vseh (tudi že umrlih) ljudeh, pri čemer naj $P(x,y)$ pomeni *"y je oče od x-a"*. V tej interpretaciji predpostavko preberemo kot *"vsak človek ima očeta"*, ki je resnična. Zaključek trdi neumnost, da *"obstaja človek, ki je oče vsem ljudem"*.

1.	$\forall x \exists y P(x,y)$	predpostavka	
2.	$\exists y P(x,y)$	US(1)	
3.	$P(x, c(x))$	ES(2)	
4.	$\forall x P(x, c(x))$	UG(3)	NAPAKA!!
5.	$\exists y \forall x P(x,y)$	EG(4)	

V vrstici 4. ne smemo uporabiti pravila UG, saj je spremenljivka x argument Skolemove funkcije $c(x)$, ki smo jo pridelali z uporabo ES v vrstici 2.

Poglavje 3

Množice

Izberimo tri naravna števila $1, 2$ in 3. Včasih bomo želeli to trojico naravnih števil shraniti v isto torbo (ali isto podatkovno strukturo, če želite) z imenom T in namesto o treh objektih razmišljati o njihovi celoti.

Takšni skupinici objektov T pravimo *množica*. Množice nam bodo omogočale, da več objektov hkrati dojemamo kot celoto.

V teoretični matematiki pojma množice tipično ne definiramo, temveč zgolj opišemo lastnosti, ki jih množice imajo. Tako lahko za objekta a in A (naivno bomo a imenovali element, A pa množica) odločimo, ali a *pripada* ali *ne pripada* množici A. Zapis

$$a \in A$$

preberemo kot "*a pripada množici A*" ali tudi "*a je element množice A*". Nasprotno možnost označimo z

$$a \notin A$$

in preberemo "*a ne pripada množici A*" ali tudi "*a ni element množice A*".

Množica je natančno določena s *svojimi elementi*, zato lahko množico podamo z naštevanjem njenih elementov (če elementov ni preveč). Tako je

$$T = \{1, 2, 3\}.$$

Isto množico T lahko zapišemo tudi kot $\{3, 2, 1\}$ ali $\{1, 1, 2, 3, 2\}$. Vrstni red naštetih elementov ni pomemben, ravno tako množica ni občutljiva na to, če katerega od njenih elementov zapišemo večkrat. Zapis $T = \{1, 1, 2, 3, 2\}$ bi lahko brali kot "1 *pripada* T *in* 1 *pripada* T *in* ..." ter poenostavili izjavo zaradi idempotence konjunkcije.

Množice smemo podajati tudi z uporabo izjavnih formul. Množico T lahko podamo tudi z opisom, da je njen element natančno vsak takšen *objekt*, ki je enak 1 ali enak 2 ali enak 3,

$$T = \{x \mid x = 1 \vee x = 2 \vee x = 3\}.$$

Zdi se, da lahko množice podamo z uporabo izjavnih formul. Denimo takole

$$A = \{x \mid \varphi(x)\}, \tag{3.1}$$

podali smo množico, ki vsebuje natančno tiste elemente x, za katere je $\varphi(x)$ resnična izjava, velja torej

$$x \in A \quad \text{natanko tedaj, ko} \quad \varphi(x). \tag{3.2}$$

Znašli smo se na nekoliko spolzkem terenu. Interpretacija izjavne formule $\varphi(x)$ je odvisna od *področja pogovora* in pomena uporabljenih predikatov. Za predikate se bomo že dogovorili, toda kaj je področje pogovora, domena?

Je to *množica* reči, ki jih lahko izbiramo za vrednost spremenljivke x? Znašli smo se v zgodbi o kuri in jajcu. Če želimo definirati množice, je potrebno imeti na voljo množico — področje pogovora.

Množice bomo obravnavali z uporabniškega stališča. Ne bo nas zanimalo, kako z uporabo aksiomatske teorije množic konstruiramo, denimo, množico naravnih ali realnih števil. Takšni dve množici za nas obstajata in pika.

Univerzalna množica S je *področje pogovora* teorije množic. Univerzalno množico bomo uporabljali kot pokrajino, iz katere smemo jemati individualne konstante (elemente), ki jih lahko vstavljamo v izjavne formule.

Od kod prihaja univerzalna množica, kateri elementi pripadajo množici S in podobna vprašanja bomo postavili na stranski tir. Mislimo si pa, da univerzalna množica S vsebuje vse elemente, ki bi jih kadarkoli potrebovali v katerikoli matematični nalogi.

Množico A, opisano z izjavno formulo $\varphi(x)$, glej (3.1), bomo dejansko opremili z vsemi tistimi elementi x, ki pripadajo univezalni množici S in zadoščajo $\varphi(x)$, oziroma simbolično

$$x \in A \quad \text{natanko tedaj, ko velja} \quad \varphi(x) \text{ in } x \in S. \tag{3.3}$$

Definirajmo še *prazno množico* $\emptyset$. Prazna množica je edina množica brez elementov, opišemo pa jo lahko z naslednjim opisom:

$$\emptyset = \{x \mid x \neq x\}.$$

Je morda pristop z univerzalno množico res potreben? Vsekakor se je pri interpretaciji izjavne formule potrebno odločiti za področje pogovora (oz. univerzalno množico).

Na tem mestu bralca opozorimo, da so lahko elementi množice tudi sámi množice. Tako lahko podamo množici

$$\{\{1,2\},\{3\}\} \quad \text{in} \quad \{\{1,2,3\}\}.$$

Pri tem velja $\{1,2\} \in \{\{1,2\},\{3\}\}$ in $\{1,2\} \notin \{\{1,2,3\}\}$.

Denimo, da množico R definiramo z naslednjim opisom

$$R = \{x \mid x \notin x\}, \tag{3.4}$$

R je množica vseh takšnih elementov, ki ne vsebujejo samega sebe kot element. Na prvi pogled z definicijo ni nič narobe, in pohlevno se smemo vprašati, ali R vsebuje samega sebe kot element, ali $R \in R$?

Z uporabo (3.1) velja

$$R \in R \text{ natanko tedaj, ko } R \notin R,$$

kar je absurdno. Še huje, ker je izjavna formula (2.14)

$$\neg\exists y \forall x (P(x,y) \Leftrightarrow \neg P(x,x))$$

splošno veljavna, množica R sploh ne obstaja.

V izjavni formuli lahko namreč za področje pogovora izberemo elemente in množice, predikat $P(x,y)$ naj pomeni pripadnost $x \in y$. V tej interpretaciji formula (2.14) kot izjava trdi, da ne obstaja (množica) y, ki za elemente vsebuje natanko tiste x-e, ki ne vsebujejo samih sebe kot element.

Z uporabo univerzalne množice opis množice R preberemo kot

$$x \in R \quad \text{natanko tedaj, ko} \quad x \notin x \text{ in } x \in S$$

S substitucijo $x = R$ pridelamo

$$\begin{aligned} R \in R \;&\Leftrightarrow R \notin R \wedge R \in S \\ &\sim \quad (R \in R \wedge R \notin R \wedge R \in S) \vee (R \notin R \wedge (R \in R \vee R \notin S) \\ &\sim \quad (R \notin R \wedge (R \in R \vee R \notin S)) \\ &\sim \quad (R \notin R \wedge R \in R) \vee (R \notin R \wedge R \notin S) \\ &\sim \quad R \notin R \wedge R \notin S \end{aligned}$$

R torej ne vsebuje same sebe kot element, ravno tako pa R ne pripada univerzalni množici.

Za konec razdelka zapišemo še dve definiciji. Množico z enim samim elementom imenujemo tudi *singleton*, množici $\{a\}$ včasih pravimo tudi *singleton* a. Množico z natanko dvema elementoma imenujemo tudi *par*.

3.1 Enakost in vsebovanost

Množici A in B sta *enaki* natanko tedaj, ko imata iste elemente. Simbolično zapišemo

$$A = B \quad \text{natanko tedaj, ko} \quad \forall x (x \in A \Leftrightarrow x \in B). \tag{3.5}$$

Množici

$$\{1,2,3\} \quad \text{in} \quad \{1,3,2\}$$

sta enaki, medtem ko množici

$$\{a, b, c, d\} \quad \text{in} \quad \{c, d, e\}$$

nista, saj črka a pripada množici $\{a, b, c, d\}$ in ne pripada množici $\{c, e, d\}$.

Za množici A in B pravimo, da je A *podmnožica* množice B, $A \subseteq B$, če za vsak element množice A velja, da pripada tudi množici B. V formalnem jeziku je

$$A \subseteq B \quad \text{natanko tedaj, ko} \quad \forall x(x \in A \Rightarrow x \in B). \tag{3.6}$$

Pravimo tudi, da množica B *vsebuje* množico A, zvezo *"biti podmnožica"* pa imenujemo tudi relacija *inkluzije*.

Množica $\{1, 2, 3\}$ je vsebovana v množici $\{1, 2, 3, 4\}$, slednja pa v množici naravnih števil $\mathbb{N}$. Velja

$$\{1, 2, 3\} \subseteq \{1, 2, 3, 4\} \subseteq \mathbb{N}.$$

Če množica A ni podmnožica množice B, bomo uporabili tudi zapis $A \not\subseteq B$.

Osnovna zveza med enakostjo in vsebovanostjo množic je naslednja.

Izrek 3.1

$$A = B \quad \textit{natanko tedaj, ko} \quad A \subseteq B \textit{ in } B \subseteq A$$

Dokaz. Računamo z uporabo definicije enakosti in vsebovanosti množic.

$$\begin{aligned} A = B \quad &\sim \quad \forall x(x \in A \Leftrightarrow x \in B) \\ &\sim \quad \forall x((x \in A \Rightarrow x \in B) \text{ in } (x \in B \Rightarrow x \in A)) \\ &\sim \quad \forall x(x \in A \Rightarrow x \in B) \text{ in } \forall x(x \in B \Rightarrow x \in A) \\ &\sim \quad A \subseteq B \text{ in } B \subseteq A \end{aligned}$$

□

Potrebovali bomo tudi relacijo *stroge vsebovanosti* ali *stroge inkluzije*. Pravimo, da je A *prava podmnožica* množice B, $A \subset B$, če je A vsebovana v B, poleg tega pa množici A in B nista enaki. Zapišemo lahko

$$A \subset B \quad \text{natanko tedaj, ko} \quad A \subseteq B \text{ in } A \neq B. \tag{3.7}$$

Izrek 3.2 *Za poljubno množico A veljajo naslednje zveze*

(i) $\emptyset \subseteq A$ *in* $A \subseteq S$

(ii) $A \subseteq A$

(iii) $A \not\subset A$

Dokaz. Za dokaz vsebovanosti $\emptyset \subseteq A$ se je potrebno prepičati, da velja

$$\forall x(x \in \emptyset \Rightarrow x \in A).$$

Izberemo poljuben x. Ker $x \notin \emptyset$, je implikacija $x \in \emptyset \Rightarrow x \in A$ resnična. Zato je resnično tudi njeno univerzalno zaprtje.

Podobno ukanemo vsebovanost $A \subseteq S$, saj za poljuben x velja $x \in S$.

Množica A je seveda enaka sama sebi, zato $A \not\subset A$. Vsebovanost $A \subseteq A$ preverimo z opazko, da je za poljuben x izjava $x \in A \Rightarrow x \in A$ resnična. □

Izrek 3.3 *Za poljubne množice A, B in C velja:*

$$\text{če} \quad A \subseteq B \quad \text{in} \quad B \subseteq C\,, \quad \text{potem} \quad A \subseteq C.$$

Dokaz. Izpeljemo ga z uporabo hipotetičnega silogizma. Zaradi obeh vsebovanosti $A \subseteq B$ in $B \subseteq C$ za poljuben x veljata implikaciji

$$x \in A \Rightarrow x \in B \quad \text{in} \quad x \in B \Rightarrow x \in C.$$

Zato za poljuben x velja tudi

$$x \in A \Rightarrow x \in C$$

in dokaz je zaključen. □

3.2 Operacije z množicami

Operacije z množicami si smemo predstavljati kot predpis[1], ki paru množic A in B priredi novo množico, rezultat operacije.

Unija množic A in B, $A \cup B$, je množica vseh elementov, ki pripadajo vsaj eni od množic A oziroma B.

$$A \cup B \quad = \quad \{x \mid x \in A \ \vee \ x \in B\} \tag{3.8}$$

Presek množic A in B, $A \cap B$, je množica vseh elementov, ki pripadajo obema množicama A in B.

$$A \cap B \quad = \quad \{x \mid x \in A \ \wedge \ x \in B\} \tag{3.9}$$

[1] Tèrmin predpis včasih uporabljamo kot ne najbolj natančen sinonim za preslikavo.

Če imata množici A in B prazen presek, $A \cap B = \emptyset$, potem pravimo, da sta *disjunktni*.

Unijo in presek množic definiramo z logičnima operacijama disjunkcijo oziroma konjunkcijo. Malenkost bolj zapleteni sta operaciji razlike in simetrične razlike množic.

Razlika množic A in B, $A \setminus B$, je množica vseh elementov, ki pripadajo A in ne pripadajo B. Razliko množic $A \setminus B$ včasih preberemo tudi kot A *brez* B.

$$A \setminus B \quad = \quad \{x \mid x \in A \ \wedge \ x \notin B\} \tag{3.10}$$

Simetrična razlika množic A in B, $A + B$, je množica vseh elementov, ki pripadajo natanko eni od množic A oziroma B.

$$A + B \quad = \quad \{x \mid x \in A \ \veebar \ x \notin B\} \tag{3.11}$$

Odkod ime simetrična razlika? Množici $A + B$ pripadajo natančno tisti elementi, ki imajo različen status pripadnosti glede na množici A in B.

Simetrično razliko definiramo z uporabo ekskluzivne disjunkcije.

Komplement množice A, označimo ga z A^c, moramo definirati z uporabo univerzalne množice S. Posamezen element pripada množici A^c natanko tedaj, ko ne pripada množici A, še vedno pa mora pripadati univerzalni množici (oz. področju pogovora) S.

$$A^c \quad = \quad \{x \mid x \notin A \ \wedge \ x \in S\}. \tag{3.12}$$

Dogovorimo se še za prednost operacij z množicami. Najmočneje veže komplement ^{c}, ki mu sledita po prednosti enakovredna presek $\cap$ in razlika $\setminus$. Unija $\cup$ in simetrična razlika $+$ sta po prednosti enakovredni operaciji, a ne vežeta tako močno kot presek ali razlika.

Prednost operacij z množicami sledi prednosti izjavnih veznikov, s katerimi so operacije definirane. Ker konjunkcija veže močneje kot disjunkcija, naj tudi presek veže močneje kot unija.

Po prednosti enakovredne operacije z množicami bomo, tako kot v izjavnem računu, združevali z leve.

^{c}	$\cap$, $\setminus$	$\cup$, $+$

Tabela 3.1: Operacije z množicami, razvrščene po prednosti.

Množico

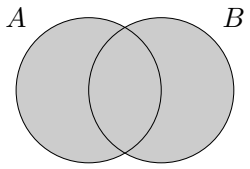
$$((((A \cap (B^c)) \setminus C) \cup D) + (E \cap F))$$

lahko v skladu z dogovorom o opuščanju oklepajev zapišemo tudi kot

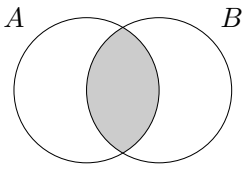
$$A \cap B^c \setminus C \cup D + E \cap F.$$

Vennovi diagrami

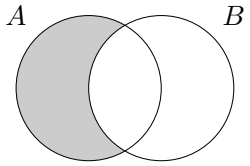
Vennovi diagrami omogočajo grafično predstavitev operacij z množicami. Množice $A, B, \ldots$ predstavimo s krogi (ali bolj zapletenimi območji) v ravnini, točke v krogu imamo za elemente posamezne množice. Osenčeno območje predstavlja rezultat operacije z množicami.



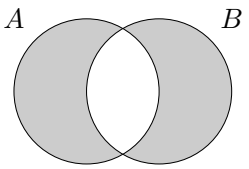
$A \cup B$



$A \cap B$

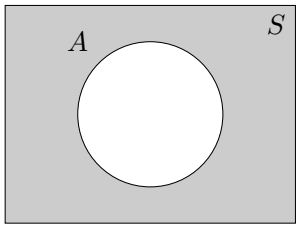


$A \setminus B$



$A + B$

Za ilustracijo komplementa množice je potrebno predstaviti univerzalno množico S. Narišemo jo lahko kot okvir, znotraj katerega rišemo kroge, ki predstavljajo posamezne množice.

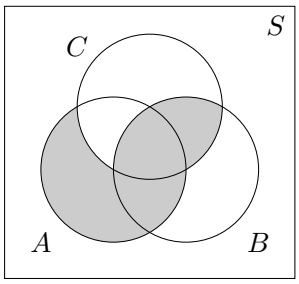


A^c

Narišimo še Vennov diagram množice

$$(A \setminus B) \cup (B \cap C).$$

Pri tem so krožnice, katerih notranjosti predstavljajo množice A, B, C, v splošni legi. Ravnino delijo na 8 območij, vsako ustreza natanko enemu *naboru logičnih vrednosti spremenljivk* p_A, p_B *in* p_C, ki označujejo, ali posamezna točka pripada notranjosti ustrezne krožnice.



$(A \setminus C) \cup (B \cap C)$

Vennove diagrame znamo sicer narisati za družine poljubno velikega števila množic. Res pa je, da z večanjem števila množic v Vennovem diagramu narašča kompleksnost območij, s katerimi posamezne množice predstavimo. Vennov diagram za tri ali štiri množice je relativno berljiv, če pa je število množic v diagramu 5 ali več, postanejo slike nepregledne.

Vsebovanost in operacije

V tem kratkem razdelku bomo spoznali nekaj zvez med relacijo podmnožice in operacijami z množicami.

Izrek 3.4 *Naj bodo A, B, C poljubne množice. Potem veljajo naslednje zveze:*

(i) *Če je $A \subseteq B$, potem je $A \cup C \subseteq B \cup C$.*

(ii) *Če je $A \subseteq B$, potem je $A \cap C \subseteq B \cap C$.*

(iii) $A \cap B \subseteq A \subseteq A \cup B$.

Dokaz. Naj bo $A \subseteq B$ in izberimo poljuben element $c \in A \cup C$. Če $c \in A$, potem $c \in B$, saj je $A \subseteq B$. Zato je $c \in B \cup C$. Če pa $c \in C$, potem c pripada tudi $B \cup C$.

Dokaz (ii) je podoben. Če $c \in A \cap C$, potem $c \in B$, saj je $A \subseteq B$. Ker $c \in C$, sledi $c \in B \cap C$.

Zveza (iii) je motivirana s poenostavitvijo in pridružitvijo, dvema praviloma sklepanja. Če velja $a \in A \ \wedge \ a \in B$, potem lahko sklepamo na $a \in A$. Podobno lahko iz $a \in A$ sklepamo na $a \in A \ \vee \ a \in B$. □

Točki (i) in (ii) izreka 3.4 pravita, da je relacija vsebovanosti *usklajena* z operacijama unije in preseka. Analogija iz aritmetike pravi, da neenakost ohranimo, če na obeh straneh neenačbe prištejemo ali odštejemo isto število. V teoriji množic pa smemo na levi in desni strani vsebovanosti z unijo (ali presekom) prilepiti isto množico, s tem vsebovanosti ne pokvarimo.

Relacijo vsebovanosti lahko enakovredno prepišemo na nekaj različnih načinov.

Izrek 3.5 *Naj bosta A in B poljubni množici. Naslednje zveze so enakovredne:*

(a) $A \subseteq B$

(b) $A \cup B = B$

(c) $A \cap B = A$

(d) $A \setminus B = \emptyset$

(e) $B^c \subseteq A^c$

Dokaz. (a) $\Rightarrow$ (b): Privzamemo, da je $A \subseteq B$. Po eni strani je $B \subseteq A \cup B$, po drugi pa z uporabo točke (i) izreka 3.4 izpeljemo zvezo $A \cup B \subseteq B \cup B$. Odtod velja $A \cup B = B$.

(a) $\Rightarrow$ (c): Z uporabo zveze $A \subseteq B$ smemo računati takole: $A = A \cap A \subseteq A \cap B \subseteq A$.

(b) $\Rightarrow$ (a): Iz $A \subseteq A \cup B$ in $A \cup B = B$ sledi $A \subseteq B$.

(c) $\Rightarrow$ (a): Tudi iz $A \cap B \subseteq B$ in $A = A \cap B$ sledi $A \subseteq B$.

(a) $\Leftrightarrow$ (d): Implikacija $x \in A \Rightarrow x \in B$ je enakovredna izrazu $x \in A \wedge x \notin B \Rightarrow x \in \emptyset$, kjer je izraz $x \in \emptyset$ kar logična konstanta 0.

(a) $\Leftrightarrow$ (e): Tokrat uporabimo kontrapozicijo, implikacija $x \in A \Rightarrow x \in B$ je enakovredna implikaciji $x \notin B \Rightarrow x \notin A$.

□

3.3 Enakosti z množicami

Operacije z množicami so definirane z uporabo logičnih operacij. Zakone izjavnega računa, enakovrednosti nekaterih parov izjavnih izrazov, lahko prepišemo v ustrezne enakosti z množicami. Konjunkcija, disjunkcija, negacija in ekskluzivna disjunkcija skoraj dobesedno ustrezajo preseku, uniji, komplementu in simetrični razliki. Tudi logični konstanti 0 in 1 na soroden način ustrezata prazni množici $\emptyset$ in univerzalni množici S.

S številčenjem sledimo številčenju zakonov izjavnega računa. Navajamo pa ponovno tudi nekatere zveze, ki veljajo za vsebovanost množic.

(1) Zakon dvojnega komplementa:

$$(A^c)^c = A$$

(2) Idempotenca:

$$A \cap A = A \qquad A \cup A = A$$

(3) Komutativnost:

$$A \cap B = B \cap A \qquad A \cup B = B \cup A$$
$$A + B = B + A$$

(4) Asociativnost:

$$(A \cap B) \cap C = A \cap (B \cap C)$$
$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A + B) + C = A + (B + C)$$

(5) Absorpcija:

$$A \cap (A \cup B) = A \qquad A \cup (A \cap B) = A$$

(6) Distributivnost:

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A + B) \cap C = (A \cap C) + (B \cap C)$$

(7) de Morganova zakona:

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

(8) Kontrapozicija:

$$A \subseteq B \ \sim \ B^c \subseteq A^c$$

(9) *Prazna množica* $\emptyset$ in *univerzalna množice* S:

$$A \cup A^c = S \qquad A \cap A^c = \emptyset$$
$$A + A = \emptyset \qquad A + A^c = S$$

(10) Substitucija $\emptyset$ in S:

$$A \cap \emptyset = \emptyset \qquad A \cup \emptyset = A$$
$$A \cap S = A \qquad A \cup S = S$$

(11) Lastnosti vsebovanosti:

$$A \subseteq B \ \sim \ A \cup B = B \ \sim \ A \cap B = A \ \sim \ A \setminus B = \emptyset$$
Če je $A \subseteq B$, potem je $A \cup C \subseteq B \cup C$.
Če je $A \subseteq B$, potem je $A \cap C \subseteq B \cap C$.
$$A \cap B \subseteq A, B \subseteq A \cup B$$

(12) Lastnosti razlike množic:

$$A \setminus B = A \cap B^c$$

(13) Lastnosti simetrične razlike:

$$A + B = (A \setminus B) \cup (B \setminus A)$$
$$A + B = (A \cup B) \setminus (A \cap B)$$

Dokazovanje zgornjih enakosti množic v kar največji meri izpustimo. Vsi dokazi sledijo naslednjemu zgledu, kjer se skličemo na ustrezen zakon izjavnega računa.

Pokažimo, da za poljubni množici A in B velja de Morganov zakon

$$(A \cup B)^c = A^c \cap B^c.$$

Izberimo poljuben element a ter poljubni množici A in B. S p_A in p_B označimo izjavni spremenljivki, ki imata isto logično vrednost kot izjavi $a \in A$ oziroma $a \in B$. Računajmo.

$$\begin{aligned} a \in (A \cup B)^c &\sim a \notin (A \cup B) \\ &\sim \neg(a \in A \cup B) \\ &\sim \neg(a \in A \vee a \in B) \\ &\sim \neg(p_A \vee p_B) \qquad (*) \\ &\sim \neg p_A \wedge \neg p_B \\ &\sim a \notin A \wedge a \notin B \\ &\sim a \in A^c \wedge a \in B^c \\ &\sim a \in A^c \cap B^c \end{aligned}$$

Edini malo netrivialni korak je v vrstici $(*)$, kjer uporabimo de Morganov zakon iz izjavnega računa.

3.4 Reševanje sistemov enačb z množicami

V tem razdelku bomo izdelali postopek za reševanje enačb z množicami z eno neznano množico. Obravnavali bomo sistem enačb z množicami

$$\begin{aligned} X \cup A &= B \setminus X \\ X \cup B &= X. \end{aligned} \tag{3.13}$$

Uporabljali bomo standardno notacijsko konvencijo, množica X naj označuje neznano množico, spremenljivko v sistemu, medtem ko so $A, B, \ldots$ množice parametri.

Rešiti sistem enačb z množicami pomeni

(S1) natančno določiti pogoje, pri katerih je sistem rešljiv in

(S2) v primeru izpolnjenih pogojev za rešljivost poiskati vse množice X, ki zadoščajo danim enačbam.

Naš prvi korak pri reševanju bo pretvorba sistema na eno samo enačbo. Dodatno bomo zahtevali, da ima ta enačba na eni strani enakosti prazno množico. Osnovni orodji sta naslednji lemi.

Lema 3.6 *$A = \emptyset$ in $B = \emptyset$ natanko tedaj, ko je $A \cup B = \emptyset$.*

Dokaz. Če je $A \cup B = \emptyset$, potem je tudi množica A prazna, saj velja $A \subseteq A \cup B = \emptyset$. Simetričen sklep velja za množico B.

V primeru, ko sta A in B prazni množici, pa velja $A \cup B = \emptyset \cup \emptyset = \emptyset$. □

Lema 3.7 *$A = B$ natanko tedaj, ko je $A + B = \emptyset$.*

Dokaz. Simetrično razliko $A + B$ lahko prepišemo kot

$$A + B \;=\; (A \setminus B) \cup (B \setminus A).$$

Če je $A + B = \emptyset$, potem lahko po lemi 3.6 sklepamo, da sta obe množici $A \setminus B$ in $B \setminus A$ prazni. To pa je enakovredno vsebovanostima $A \subseteq B$ in $B \subseteq A$, zato je $A = B$.

Če sta množici A in B enaki, je seveda njuna simetrična razlika $A + B$ prazna. □

Prav na koncu postopka bomo obravnavali enačbo oblike $(X \cap P) \cup (X^c \cap Q) = \emptyset$. Razdelajmo, kdaj je rešljiva takšna enakost.

Lema 3.8 *Enačba*

$$(X \cap P) \cup (X^c \cap Q) = \emptyset \tag{3.14}$$

je rešljiva natanko tedaj, ko velja

$$Q \subseteq P^c.$$

V tem primeru so rešitve natanko vse množice X, za katere velja

$$Q \subseteq X \subseteq P^c. \tag{3.15}$$

Dokaz. Če za množico X velja $Q \subseteq X$, potem je $X^c \cap Q = Q \setminus X = \emptyset$. Če velja tudi $X \subseteq P^c$, potem je $P \subseteq X^c$ in zato je $P \setminus X^c = P \cap (X^c)^c = P \cap X = \emptyset$. Vsaka množica X, za katero velja $Q \subseteq X \subseteq P^c$, je torej rešitev enačbe (3.14).

Kdaj takšna množica X obstaja? Natanko tedaj, ko velja pogoj $Q \subseteq P^c$, ki je torej zadostni pogoj za rešljivost enačbe (3.14).

Za drugo smer razmišljamo takole. Naj bo množica X rešitev enačbe (3.14). Po lemi 3.6 sklepamo, da sta obe množici $X \cap P$ in $X^c \cap Q$ prazni. Iz $X \cap P = \emptyset$ sledi, da je $X^c \cap P = P$. Zato je $P \subseteq X^c$, kar enakovredno prepišemo v zvezo $X \subseteq P^c$. Iz $X^c \cap Q = \emptyset$ pa sledi vsebovanost $Q \subseteq X$.

Torej za vsako rešitev X enačbe (3.14) velja $X \subseteq P^c$ in $Q \subseteq X$. Oziroma, če je ta enačba rešljiva in je X poljubna rešitev, mora veljati $Q \subseteq P^c$ in tudi $Q \subseteq X \subseteq P^c$. □

Še dodatni komentar. Poljubno množico X, ki zadošča pogoju (3.15), lahko v parametrični obliki zapišemo z enačbo

$$X = Q \cup (T \cap P^c),$$

pri čemer je T *res čisto* poljubna množica.

Vrnimo se k začetni lemi 3.7. Omogoča nam, da sistem (3.13) prepišemo z enačbami, ki imajo na desni strani prazno množico.

$$\begin{aligned} (X \cup A) + (B \setminus X) &= \emptyset \\ (X \cup B) + X &= \emptyset \end{aligned}$$

Z uporabo leme 3.6 ga lahko prepišemo v eno samo enakost.

$$((X \cup A) + (B \setminus X)) \cup ((X \cup B) + X) = \emptyset \tag{3.16}$$

Pri tem je enačba (3.16) enakovredna prvotnemu sistemu enačb (3.13).

Računajmo:

$$\begin{aligned} \emptyset &= ((X \cup A) + (B \setminus X)) \cup ((X \cup B) + X) = \\ &= ((X \cup A) \setminus (B \setminus X)) \cup ((B \setminus X) \setminus (X \cup A)) \cup ((X \cup B) \setminus X) \cup (X \setminus (X \cup B)) \\ &= ((X \cup A) \cap (B \cap X^c)^c) \cup ((B \cap X^c) \cap (X \cup A)^c) \cup ((X \cup B) \cap X^c) \cup (X \cap (X \cup B)^c) \\ &= ((X \cup A) \cap (B^c \cup X)) \cup ((B \cap X^c) \cap (X^c \cap A^c)) \cup ((X \cup B) \cap X^c) \cup (X \cap (X^c \cap B^c)) \\ &= (X \cap B^c) \cup (X \cap X) \cup (A \cap B^c) \cup (A \cap X) \cup \\ &\qquad \cup (B \cap X^c \cap X^c \cap A^c) \cup (X \cap X^c) \cup (B \cap X^c) \cup (X \cap X^c \cap B^c) \end{aligned}$$

Členi v zadnjem izrazu so vezani z unijami. Po zgradbi pa so posamezni členi preseki množic in/ali njihovih komplementov. Člene ločimo glede na uporabo neznane množice X oziroma njenega komplementa X^c.

Če člen vsebuje presek $X \cap X^c$, potem je, ne glede na to, kakšna je množica X, prazen. Zato se v izrazu lahko znebimo, denimo, člena $X \cap X^c \cap B^c$.

Če v členu W neznana množica ne nastopa, lahko tak člen zaradi enakosti $W = S \cap W = (X \cup X^c) \cap W = (X \cap W) \cup (X^c \cap W)$ nadomestimo z unijo $(X \cap W) \cup (X^c \cap W)$. V našem primeru torej člen $A \cap B^c$ nadomestimo z unijo $(X \cap A \cap B^c) \cup (X^c \cap A \cap B^c)$.

S to operacijo dosežemo, da vsak člen vsebuje presek z natančno eno izmed množic X oziroma X^c. Lahko jih uredimo.

$$\begin{aligned} &= (X \cap B^c) \cup (X) \cup (X \cap A \cap B^c) \cup (X \cap A) \cup \\ &\quad \cup (X^c \cap A \cap B^c) \cup (X^c \cap A^c \cap B) \cup (X^c \cap B) \end{aligned}$$

Na tem mestu upoštevamo absorpcijo. Če za člena velja $W \subseteq W'$, lahko člen W iz unije izpustimo in račun nadaljujemo.

$$\begin{aligned} &= (X) \cup (X^c \cap A \cap B^c) \cup (X^c \cap B) \\ &= (X) \cup (X^c \cap ((A \cap B^c) \cup B)) \\ &= (X) \cup (X^c \cap ((A \cup B) \cap (B^c \cup B))) \\ &= (X) \cup (X^c \cap (A \cup B)) \end{aligned} \tag{3.17}$$

Z zgodbo smo pri koncu. Če upoštevamo sam začetek računa, smo pridelali enačbo oblike, ki jo zahteva lema 3.8. Pri tem za množico P izberemo univerzalno množico S in za Q množico $A \cup B$. Sistem (3.13) je torej rešljiv natanko tedaj, ko je izpolnjen pogoj

$$A \cup B \subseteq \emptyset, \tag{3.18}$$

pri čemer so rešitve natanko tiste množice X, ki zadoščajo zvezi

$$A \cup B \subseteq X \subseteq \emptyset. \tag{3.19}$$

Z drugimi besedami. Če naj bo sistem (3.13) rešljiv, morata tako A kot B biti prazni množici, edina rešitev pa je $X = \emptyset$.

Opisani postopek reševanja sistemov pripelje do uspeha. V nekaterih primerih pa lahko sistem rešimo tudi precej hitreje.

Iz prve enačbe sistema (3.13) lahko pridelamo zvezo

$$X \subseteq X \cup A = B \setminus X = B \cap X^c \subseteq X^c.$$

Vsaka rešitev omenjenega sistema X mora biti vsebovana v svojem komplementu. To je možno samo v primeru $X = \emptyset$, kar je edina možnost za rešitev. Vstavimo torej $X = \emptyset$ v sistem (3.13).

$$\begin{aligned} A = \emptyset \cup A &= B \setminus \emptyset = B \\ B = \emptyset \cup B &= \emptyset \end{aligned}$$

Kdaj sta enačbi izpolnjeni? Iz druge enakosti sledi, da je $B = \emptyset$. Iz prve vrstice pa pridelamo, da velja $A = B$, posledično je tudi $A = \emptyset$.

Pridelali smo natanko isti pogoj rešljivosti, $A = B = \emptyset$, in tudi isto rešitev $X = \emptyset$.

Prvi postopek se zdi dolg in potraten, četudi poteka direktno. Druga možnost, ne da se je opisati kot postopek, je precej učinkovitejša. Morda gre enostavnost alternativne

možnosti tudi na rovaš enostavne rešitve in enostavnega pogoja rešljivosti. V primeru, ko sta pogoj za rešljivost in struktura rešitve bolj zapletena, nas ad-hoc metoda ne bo privedla do končnega rezultata.

Na koncu naj omenimo, da smemo v sistemih dopustiti tudi vrstice, ki ne predstavljajo enakosti, temveč vsebovanosti. Zvezo $L \subseteq R$ znamo namreč enakovredno prepisati kot enakost $L \cup R = R$.

3.5 Družine množic

Tudi množice lahko nastopajo kot elementi v množicah. Množica sicer ne more pripadati sama sebi, vseeno pa dopuščamo, da je množica element kakšne druge množice.

V tem razdelku si bomo ogledali nekaj primerov množic, ki imajo za elemente množice. Takšnim konstruktom bi lahko rekli *množice množic*, pa bomo vseeno raje uporabljali jeziku prijaznejše ime *družina množic*.

Potenčna množica

Oglejmo si množico

$$A = \{1, 2, 3\}$$

in se vprašajmo, katere podmnožice množice A lahko zapišemo. Najprej je tu prazna množica $\emptyset$, potem tri množice z enim samim in tri s po dvema elementoma, nazadnje pa tudi množica A. Tudi zanjo namreč velja $A \subseteq A$.

Potenčna množica množice A, $\mathcal{P}A$, je družina vseh podmnožic množice A. Velja torej

$$\mathcal{P}A = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Za vsakega od elementov $1, 2, 3$ iz A se lahko odločimo, ali ga dopustimo v podmnožici ali ne. Takšno izbiro elementov v podmnožici lahko predstavimo s trojico logičnih vrednosti. Vseh podmnožic množice A je torej natančno toliko, kot je naborov logičnih vrednosti za tri izjavne spremenljivke. V našem zgledu 2^3.

Isti premislek deluje tudi za množice z drugim številom elementov. Kar brez dokaza lahko zapišemo naslednji izrek.

Izrek 3.9 *Če je A množica z $n \in \mathbb{N}$ elementi, potem ima njena potenčna množica $\mathcal{P}A$ natanko 2^n elementov, ali enakovredno, A ima natanko 2^n podmnožic.*

Kakšna pa je potenčna množica prazne množice $\mathcal{P}\emptyset$? Prazna množica ima natanko 0 elementov, njena potenčna množica naj bi jih imela natančno $2^0 = 1$. In res, $\emptyset$ je edina podmnožica prazne množice, zato je

$$\mathcal{P}\emptyset = \{\emptyset\}.$$

Izračunajmo še $\mathcal{P}\mathcal{P}\emptyset = \mathcal{P}\{\emptyset\}$. Množica $\{\emptyset\}$ ima natančno en element, v njeni potenčni množici pričakujemo natanko $2^1 = 2$ elementa. Res sta $\emptyset$ in $\{\emptyset\}$ edini podmnožici množice $\{\emptyset\}$, zato je

$$\mathcal{P}\mathcal{P}\emptyset = \mathcal{P}\{\emptyset\} = \{\emptyset, \{\emptyset\}\}.$$

3.5.1 Unija in presek družine

Izberimo množico $\mathcal{I}$, ki jo imenujmo *indeksna množica*. *Družina množic* je preslikava[2] $\mathcal{A}$, ki vsakemu indeksu $i \in \mathcal{I}$ določi množico A_i. Simbolično pišemo tudi

$$\mathcal{A} = \{A_i \mid i \in \mathcal{I}\}.$$

Množice $A_i, i \in \mathcal{I}$, imenujemo tudi *člane* družine $\mathcal{A}$.

Unija družine $\mathcal{A}$ je množica vseh elementov, ki pripadajo vsaj enemu bloku družine $\mathcal{A}$. Simbolično

$$\bigcup \mathcal{A} = \bigcup_{i \in \mathcal{I}} A_i = \{x \mid \exists i (i \in \mathcal{I} \wedge x \in A_i)\}. \tag{3.20}$$

Analogno lahko definiramo *presek družine* $\mathcal{A}$ kot množico tistih elementov, ki pripadajo vsem blokom družine. Zapišemo lahko

$$\bigcap \mathcal{A} = \bigcap_{i \in \mathcal{I}} A_i = \{x \mid \forall i (i \in \mathcal{I} \wedge x \in A_i)\}. \tag{3.21}$$

Definirajmo družino $\mathcal{A}$ z indeksno množico $\mathbb{N} \setminus \{0, 1\} = \{2, 3, \ldots\}$ in opisom

$$A_i \text{ je odprti interval realnih števil med } \frac{1}{i} \text{ in } i, \qquad A_i = (1/i, i).$$

Potem je

$$\bigcup \mathcal{A} = \bigcup_{i=2}^{\infty} A_i = (0, \infty) \quad \text{in}$$
$$\bigcap \mathcal{A} = \bigcap_{i=2}^{\infty} A_i = A_2 = (1/2, 2).$$

Če je indeksna množica $\mathcal{I}$ končna, denimo $\mathcal{I} = \{1, 2, 3, 4\}$, potem sta unija in presek družine $\mathcal{A}$ identična uniji in preseku blokov družine,

$$\bigcup \mathcal{A} = \bigcup_{i=1}^{4} A_i = A_1 \cup A_2 \cup A_3 \cup A_4 \quad \text{in}$$
$$\bigcap \mathcal{A} = \bigcap_{i=1}^{4} A_i = A_1 \cap A_2 \cap A_3 \cap A_4.$$

[2] Preslikave definiramo v poglavju 5.

Pravimo, da je družina množic $\mathcal{A} = \{A_i \mid i \in \mathcal{I}\}$ *pokritje*[3] množice B, če je

(P1) $\bigcup \mathcal{A} = \bigcup_{i \in \mathcal{I}} A_i = B$.

Za družino množic $\mathcal{A} = \{A_i \mid i \in \mathcal{I}\}$ pravimo, da je *razbitje* množice B, če velja

(R1) $\mathcal{A}$ je pokritje množice B, $\bigcup \mathcal{A} = \bigcup_{i \in \mathcal{I}} A_i = B$,

(R2) bloki družine $\mathcal{A}$ so neprazni, za vsak indeks $i \in \mathcal{I}$ velja $A_i \neq \emptyset$, in

(R3) bloki družine so paroma disjunktni, za vsaka različna indeksa $i, j \in \mathcal{I}$ velja $A_i \cap A_j = \emptyset$.

Kot zgled znova ponudimo intervale na realni osi, ki jih indeksiramo s celimi števili. Družina $\mathcal{C} = \{[i, i+1] \mid i \in \mathbb{Z}\}$ je pokritje množice realnih števil $\mathbb{R}$. Ni pa razbitje, saj, denimo, število 11 pripada tako intervalu $[10, 11]$ kot tudi $[11, 12]$.

Družina $\mathcal{H} = \{[i, i+1) \mid i \in \mathbb{Z}\}$ je razbitje množice realnih števil $\mathbb{R}$.

3.6 Kartezični produkt množic

Urejene pare želimo definirati kategorično — ne zanima nas, kaj urejeni par je, temveč kako se obnaša. Tako je *urejen par* (a, b) s *prvo koordinato* a in *drugo koordinato* b struktura, za katero velja t.i. *osnovna lastnost urejenih parov*:

$$(a, b) = (c, d) \text{ natanko tedaj, ko je } a = c \text{ in } b = d.$$

Urejeni par si smemo predstavljati kot vrečko, v katero shranimo dva objekta, vrečka sama pa zna shranjena objekta ločiti — eden je na prvem, drugi na drugem mestu.

Množica sama te sposobnosti nima, ne zmore ločiti med svojima elementoma in enega od njiju postaviti pred drugega. Presenetljivo je, da *lahko* urejene pare definiramo zgolj z uporabo množic.

Trditev 3.10 *Denimo, da urejen par* (a, b) *definiramo kot* $\{\{a\}, \{a, b\}\}$. *Za tako definirane urejene pare velja osnovna lastnost.*

Dokaz. Pokazati je potrebno, da velja ekvivalenca

$$\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\} \quad \text{natanko tedaj, ko je} \quad a = c \text{ in } b = d.$$

Implikacija v levo je poceni. Če veljata enakosti $a = c$ in $b = d$, sta seveda množici na levi strani ekvivalence enaki.

Za implikacijo v desno bo potrebno obravnavati dva primera, $a = b$ in $a \neq b$. Če je $a \neq b$, potem $\{a\} \neq \{a, b\}$. Množici na levi sta enaki in vsebujeta iste elemente, zato

[3] Včasih za pokritje zahtevamo le $B \subseteq \bigcup \mathcal{A} = \bigcup_{i \in \mathcal{I}} A_i$.

$\{a\} \in \{\{c\}, \{c,d\}\}$ in $\{a,b\} \in \{\{c\}, \{c,d\}\}$. Ker $\{a,b\} \neq \{c\}$, velja $\{a,b\} = \{c,d\}$ in zato $\{a\} = \{c\}$. Zato je tudi $a = c$ in posledično $b = d$.

Če pa je $a = b$, potem je

$$\{\{a\}, \{a,b\}\} = \{\{a\}, \{a,a\}\} = \{\{a\}, \{a\}\} = \{\{a\}\}.$$

Odtod sklepamo, da je $\{a\}$ edini element množice $\{\{c\}, \{c,d\}\}$. Zato velja $\{a\} = \{c\} = \{c,d\}$ in odtod sklepamo na $a = c$ in $c = d$. Ker po predpostavki velja $a = b$, lahko izpeljemo tudi enakost $b = d$. □

Kartezični produkt množic A in B, $A \times B$, je množica vseh urejenih parov s prvo koordinato iz A in drugo koordinato iz B,

$$A \times B = \{(a,b) \mid a \in A \text{ in } b \in B\}. \tag{3.22}$$

Namesto urejenih parov lahko definiramo tudi urejene trojice, četverke, in tudi urejene n-terice. Pri tem ni pomembno kako jih definiramo, pomembno je, da zadoščajo osnovni lastnosti — dve n-terici sta enaki natanko tedaj, ko se ujemata v vseh koordinatah.

Definicijo kartezičnega produkta razširimo na več faktorjev. Tako je $A \times B \times C$ množica urejenih trojic s prvo koordinato iz A, drugo iz B in tretjo iz množice C.

Kakšne so lastnosti kartezičnega produkta? Iz samega opisa urejenih parov sledi, da kartezični produkt ni komutativen. V splošnem namreč velja $A \times B \neq B \times A$, saj, denimo, urejena para (a,b) in (b,a) nista enaka.

(1) Prazen kartezični produkt:

$$A \times B = \emptyset \quad \text{natanko tedaj, ko je} \quad A = \emptyset \quad \text{ali} \quad B = \emptyset.$$

(2) Kartezični produkt in unija:

$$A \times (B \cup C) = (A \times B) \cup (A \times C),$$
$$(A \cup B) \times C = (A \times C) \cup (B \times C).$$

(3) Kartezični produkt in presek:

$$(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D) = (A \times D) \cap (B \times C).$$

(4) Monotonost kartezičnega produkta:

$$\text{Če je } A \subseteq C \quad \text{in} \quad B \subseteq D\,, \quad \text{potem je} \quad A \times B \subseteq C \times D.$$

(5) Monotonost v obratni smeri:

$$\text{Če je } A \times B \subseteq C \times D \quad \text{in} \quad A \times B \neq \emptyset\,, \quad \text{potem je} \quad A \subseteq C \wedge B \subseteq D.$$

Dokaze zgornjih lastnosti kartezičnega produkta je najlaže zapisati z uporabo definicij operacij z množicami. Tako velja

$$\begin{aligned}(x,y) \in A \times (B \cup C) \quad &\sim \quad x \in A \text{ in } y \in B \cup C \\ &\sim \quad x \in A \text{ in } (y \in B \text{ ali } y \in C) \\ &\sim \quad (x \in A \text{ in } y \in B) \text{ ali } (x \in A \text{ in } y \in C) \\ &\sim \quad (x,y) \in A \times B \text{ ali } (x,y) \in A \times C \\ &\sim \quad (x,y) \in (A \times B) \cup (A \times C).\end{aligned}$$

Zato velja $A \times (B \cup C) = (A \times B) \cup (A \times C)$.

Analogne dokaze preostalih zvez izpustimo.

Vseeno pa primerjajmo zvezi med unijo in kartezičnim produktom ter presekom in kartezičnim produktom. Zvezi (2) veljata tudi v variantah, zapisanih s presekom. Če namreč v enakosti (3) vstavimo možnost $C = D$, velja zveza

$$\begin{aligned}(A \cap B) \times C &= (A \cap B) \times (C \cap C) \\ &= (A \times C) \cap (B \times C).\end{aligned}$$

Enakost (3), prepisana z unijami namesto preseki, pa ni veljavna. Protiprimer lahko konstruiramo z uporabo lastnosti (1). V primeru, ko sta množici B in C prazni, je tudi $(A \times C) \cup (B \times D)$ prazna množica. Z izbiro poljubnih nepraznih množic A in D pa konstruiramo neprazen kartezični produkt $(A \cup B) \times (C \cup D)$.

Ravno tako lahko imamo (4) in (5) za skoraj nasprotni implikaciji. Pokažimo, da je predpostavka $A \times B \neq \emptyset$ nujno potrebna, če želimo iz vsebovanosti kartezičnih produktov

$$A \times B \subseteq C \times D$$

izpeljati vsebovanosti posameznih faktorjev $A \subseteq C$ in $B \subseteq D$.

Če želimo pokazati, da je $A \subseteq C$, je potrebno za poljuben element $a \in A$ pokazati, da pripada tudi množici C. Če je množica $A \times B$ neprazna, potem je zaradi (1) neprazna tudi množica B in obstaja $b' \in B$. Zato tudi $(a, b') \in A \times B \subseteq C \times D$ in posledično $a \in C$.

Če je kartezični produkt $A \times B$ prazen, parov s prvo koordinato a v $A \times B$ sploh ni. Četudi je za vsak urejen par (x, y) pravilna implikacija

$$(x,y) \in A \times B \;\Rightarrow\; (x,y) \in C \times D,$$

ne moremo sklepati, da je resničen tudi njen konsekvens. Protiprimer je naslednja izbira množic

$$A = \{a, a'\}\,,\; B = \emptyset\,,\; C = \{a'\}\,,\; D = \{d\}.$$

Poglavje 4

Relacije

Izberimo množico A in naj nam služi tudi kot področje pogovora. Izberimo še enomestni predikat P na področju pogovora A. Za vsak $a \in A$ je izjava $P(a)$ bodisi resnica ali laž in

$$\{a \mid a \in A \text{ in } P(a)\}$$

je podmnožica množice A, *opisana* s predikatom P.

Zgodbo lahko povemo tudi v drugi smeri. Naj bo $A' \subseteq A$ in z njeno pomočjo definirajmo enomestni predikat P' z opisom

$$P'(a) \text{ je resnična izjava natanko tedaj, ko } a \in A'.$$

V dvomestni predikat, definiran na področju pogovora A, vstavljamo urejene pare elementov množice A. Množica urejenih parov, za katere je tak dvomestni predikat resnična izjava, je *podmnožica* kartezičnega produkta $A \times A$.

Obravnava dvomestnih predikatov s stališča teorije množic (in ne matematične logike) nas napelje k naslednji definiciji.

R je *(dvomestna) relacija* v množici A, če je R podmnožica kartezičnega produkta $A \times A$, $R \subseteq A \times A$.

Definiramo lahko tudi večmestne relacije kot alternativni pogled na večmestne predikate.

Vseeno se bomo v tem poglavju ukvarjali zgolj z dvomestnimi relacijami. Zgodba bo dovolj zanimiva že v tem primeru.

Naj bo R relacija v množici A. Dejstvo, da $(a, b) \in R$, v disciplini relacij ponavadi zapišemo v *vmesni* (*infiksni*) obliki kot

$$aRb \tag{4.1}$$

in beremo *"a je v relaciji R z b."*.

Oglejmo si nekaj zgledov (dvomestnih) relacij.

- Relacija $R = \{(a,b),(b,c),(c,d),(c,c)\}$ v množici $A = \{a,b,c,d\}$.
- Relacija *manjše ali enako* $\leq$ v množici naravnih števil $\mathbb{N}$. Relacija $\leq$ je enaka

  $$\{(0,0),(0,1),(0,2),\ldots,(1,1),(1,2),\ldots,(3,5),\ldots\}.$$

  Vmesni način zapisa $3 \leq 5$ je precej bolj naraven kot zapis $(3,5) \in \leq$. Slednji se nam zdi tudi sintaktično sumljiv.

  Sorodne so relacije številskih urejenosti $<, >, \geq$, tudi v drugih številskih množicah $\mathbb{Z}, \mathbb{Q}$ in $\mathbb{R}$.
- V družini izjavnih izrazov $\mathcal{I}$ poznamo relacijo enakovrednosti $\sim$. Ravno tako lahko v isti množici definiramo relacijo "*logično sledi iz*" z opisom

  $$I_1 \text{ logično sledi iz } I_2 \text{ natanko tedaj, ko } I_2 \models I_1.$$
- Relacija *vsebovanosti* $\subseteq$ v družini podmnožic izbrane množice A.
- *Univerzalna relacija* U_A in *prazna relacija* $\emptyset$ v množici A. Univerzalna relacija U_A je kar enaka kartezičnemu produktu $A \times A$. Tudi prazna množica $\emptyset$ je podmnožica $A \times A$. Če jo imamo za relacijo v množici A, jo bomo imenovali *prazna relacija*.
- *Relacija identitete* ali *enakosti* id_A v množici A je množica vseh urejenih parov z enakima koordinatama,

  $$\mathrm{id}_A = \{(a,a) \mid a \in A\}.$$
- V množici ljudi lahko definiramo cel spekter različnih sorodstvenih relacij (zvez). Kot primera predstavimo relaciji "*mož*" in "*hči*", definirani z opisoma

  $$x \textit{ mož } y \text{ natanko tedaj, ko je } x \text{ mož od } y\text{-a.}$$
  $$x \textit{ hči } y \text{ natanko tedaj, ko je } x \text{ hči od } y\text{-a.}$$

Naj bo R relacija v množici A. *Definicijsko območje* relacije R, $\mathcal{D}_R$, je množica vseh prvih koordinat parov iz relacije R. *Zaloga vrednosti* relacije R, $\mathcal{Z}_R$, je množica vseh drugih koordinat parov iz relacije R.

$$\mathcal{D}_R = \{x \mid \exists y \in A(xRy)\}$$
$$\mathcal{Z}_R = \{y \mid \exists x \in A(xRy)\}$$

Za relacijo $R = \{(a,b),(b,c),(c,d),(c,c)\}$ je njeno definicijsko območje $\mathcal{D} = \{a,b,c\}$, zaloga vrednost pa je enaka $\mathcal{Z}_R = \{b,c,d\}$.

Relacija "*hči*" ima definicijsko območje enako množici vseh oseb ženskega spola, vsaka ženska/punca/deklica je svojih staršev hči. Zaloga vrednosti relacije "*hči*" pa je množica vseh tistih ljudi, ki imajo (vsaj eno) hčer.

Za konec razdelka definirajmo še zožitev relacije na podmnožico. Naj bo R relacija v množici A in $B \subseteq A$. *Zožitev relacije* R na množico B definiramo kot

$$R \cap (B \times B).$$

4.1 Lastnosti relacij

V tem razdelku bomo opisali nekaj lastnosti relacij. Relacije, ki imajo podobne lastnosti, bomo želeli obravnavati hkrati. V nadaljevanju bomo tako hkrati obravnavali ekvivalenčne relacije, poseben razdelek pa bo namenjen tudi relacijam urejenosti.

Naj bo R relacija v množici A. Pravimo, da je relacija

- R *refleksivna* natanko tedaj, ko $\forall x\, xRx$,
- R *simetrična* natanko tedaj, ko $\forall x \forall y\, (xRy \Rightarrow yRx)$,
- R *antisimetrična* natanko tedaj, ko $\forall x \forall y\, (xRy \wedge yRx \Rightarrow x = y)$,
- R *tranzitivna* natanko tedaj, ko $\forall x \forall y \forall z\, (xRy \wedge yRz \Rightarrow xRz)$,
- R *sovisna* natanko tedaj, ko $\forall x \forall y\, (x \neq y \Rightarrow xRy \vee yRx)$,
- R *enolična* natanko tedaj, ko $\forall x \forall y \forall z\, (xRy \wedge xRz \Rightarrow y = z)$.

Relacija je refleksivna, če je vsak element v relaciji sam s sabo. Refleksivne so univerzalna relacija, relacija identitete, pa tudi relacija vsebovanosti $\subseteq$ in relaciji $\leq$ in $\geq$ na številskih množicah $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$. Relacija stroge vsebovanosti $\subset$ in relaciji strogih neenakosti $<$ in $>$ nista refleksivni. Ravno tako ni refleksivna relacija "*hči*", saj nihče ni sam svoja hči.

Relacija R je simetrična, če iz aRb sledi bRa. Univerzalna relacija in relacija enakosti sta simetrični, ravno tako je simetrična tudi relacija vzporednosti $\parallel$ v množici premic v ravnini (ali prostoru). V množici ljudi sta simetrični denimo relaciji "*zakonec*" in "*sorodnik*".

Tipične antisimetrične relacije so $\subseteq$, $\leq$ in $\geq$. Na trivialen način, ker antecedens implikacije $aRb \wedge bRa$ ni nikoli izpolnjen, pa so antisimetrične tudi stroge neenakosti $<$ in $>$, v množici ljudi pa denimo relacija "*hči*".

Tranzitivnost relacije bomo tipično povezovali z urejenostjo. Implikacija

$$aRb \wedge bRc \Rightarrow aRc$$

lahko pomeni, da bo *a prej na vrsti na blagajni kot c*, če je a v vrsti pred b-jem in b v vrsti pred c-jem. Tranzitivne so relacije $\subseteq$, $\leq$, $<$, $\geq$, $>$, pa tudi univerzalna relacija in relacija enakosti. Relacija "*hči*" in relacija "*sorodnik*"[1] pa denimo nista tranzitivni relaciji.

Sovisne so denimo relacije številskih urejenosti $\leq$, $<$, $\geq$, $>$, relacija vsebovanosti $\subseteq$ pa v splošnem ni sovisna. Množici A in B bosta glede na relacijo vsebovanosti $\subseteq$ tipično neprimerljivi.

[1] Sorodnikov sorodnik je morda z nami le v *svaštvu*.

Relacija R je enolična v množici A, če za vsak element $a \in A$ velja, da je v relaciji R s kvečjemu enim elementom. Natančneje, v relaciji R obstaja največ en urejen par s prvo koordinato enako a. Enolična je na primer relacija enakosti, pa tudi, če se držimo črke zakona, relacija "*zakonec*". Relaciji "*hči*" in "*oče*" nista enolični. Vsaka ženska je hči tako svoje matere kot tudi očeta, po drugi strani pa ima lahko moški več otrok, torej je lahko oče različnim ljudem.

4.2 Operacije z relacijami

Seštevanje v množici naravnih števil in unija množic sta tipična zgleda operacij. Dve naravni števili lahko seštejemo, ravno tako lahko konstruiramo unijo dveh množic. Dobljeni rezultat je znova naravno število oziroma množica.

Operacije z relacijami bomo želeli definirati v takšnem, morda še malenkost bolj omejujočem smislu. Če sta R in S relaciji v množici A, bomo za rezultat operacije (ki upošteva relaciji R in S) znova želeli relacijo v *isti* množici A.

Operacije z množicami, unijo $\cup$, presek $\cap$, razliko $\setminus$ in simetrično razliko $+$, lahko v domeno relacij prestavimo brez težav. Če sta R in S relaciji v množici A, potem so takšne tudi

$$R \cup S, \quad R \cap S, \quad R \setminus S \quad \text{in} \quad R + S.$$

Vse omenjene relacije so podmnožice kartezičnega produkta $A \times A$.

Komplement relacije R je smiselno definirati kot množico vseh parov kartezičnega produkta $A \times A$, ki niso v relaciji R. Velja torej

$$R^c = \{(x, y) \mid (x, y) \in A \times A \ \text{ in } \ \neg xRy\},$$

kar je enakovredno razliki relacije R z univerzalno relacijo U_A,

$$R^c = U_A \setminus R = (A \times A) \setminus R.$$

Unija relacij $<$ in $\mathrm{id}_{\mathbb{N}}$ v množici naravnih števil $\mathbb{N}$ je relacija $\leq$ v isti množici. Ker gre za disjunktno unijo, relaciji $<$ in $\mathrm{id}_{\mathbb{N}}$ sta namreč disjunktni, velja tudi, da je relacija $<$ razlika relacij $\leq$ in $\mathrm{id}_{\mathbb{N}}$.

Presek relacij $\leq$ in $\geq$ je relacija enakosti $\mathrm{id}_{\mathbb{N}}$. Komplement relacije $<$ je relacija $\geq$. Za vsak par naravnih števil a, b je resnična natanko ena od možnosti $a < b$ oziroma $a \geq b$.

Struktura relacij dopušča definicijo dodatnih dveh operacij.

Inverzno relacijo k relaciji R, označimo jo z R^{-1}, dobimo tako, da zamenjamo koordinati v vseh parih relacije R. Velja namreč

$$R^{-1} = \{(y, x) \mid (x, y) \in R\}$$

ali enakovredno

$$xR^{-1}y \quad \text{natanko tedaj, ko je} \quad yRx.$$

Inverzna relacija k relaciji $\leq$ je, ne presenetljivo, relacija $\geq$. Za (naravni) števili a in b velja $a \leq b$ natanko tedaj, ko je $b \geq a$. Na področju sorodstvenih relacij med ljudmi pa sta si medsebojno inverzni relaciji "*otrok*" in "*roditelj*", medtem ko je relacija "*zakonec*" inverzna sama sebi.

Produkt relacij $R * S$ definiramo z naslednjim opisom

$$xR * Sy \quad \text{natanko tedaj, ko} \quad \exists z(xRz \text{ in } zSy). \tag{4.2}$$

Za zgled izberimo produkt relacije "*mož*" in relacije "*hči*" v množici ljudi. Po definiciji produkta je

$$x \ mož * hči \ y \quad \text{natanko tedaj, ko} \quad \exists z(x \ mož \ z \text{ in } z \ hči \ y),$$

tj. ko obstaja človek z, za katerega velja, da je x *mož od* z-ja in je z *hči* od osebe y. Torej je x *mož od hčere od* y, ali ekvivalentno x je *zet* od y.

Zato je produkt relacij "*mož* $*$ *hči*" enak relaciji "*zet*". Hkrati ugotovimo, da produkt relacij v splošnem ni komutativen, saj relaciji "*mož* $*$ *hči*" ter "*hči* $*$ *mož*" nista enaki.

V nadaljevanju naštejmo nekaj enakosti z relacijami, ki se tičejo operacij inverzne relacije in relacijskega produkta.

Trditev 4.1 *Naj bodo R, S, T poljubne relacije v množici A. Potem veljajo naslednje enakosti:*

(i) $(R^{-1})^{-1} = R$ *(zakon dvojnega inverza)*

(ii) $(R * S)^{-1} = S^{-1} * R^{-1}$

(iii) $(R * S) * T = R * (S * T)$ *(asociativnost produkta relacij)*

(iv) $R * (S \cup T) = (R * S) \cup (R * T)$ *in* $(R \cup S) * T = (R * T) \cup (S * T)$
(distributivnost produkta in unije)

(v) $R * \mathrm{id}_A = \mathrm{id}_A * R = R$ *in* $R * \emptyset = \emptyset * R = \emptyset$

(vi) *Če velja* $R \subseteq S$, *potem sledi* $R * T \subseteq S * T$ *in* $T * R \subseteq T * S$.

Dokaz. Zakon dvojnega inverza (i) sledi zakonom dvojne negacije in dvojnega komplementa, utemeljitev izpustimo.

Pri (ii) računamo takole:

$$\begin{aligned} a(R * S)^{-1}b &\sim bR * Sa \\ &\sim \exists x\,(bRx \text{ in } xSa) \\ &\sim \exists x\,(aS^{-1}x \text{ in } xR^{-1}b) \\ &\sim aS^{-1} * R^{-1}b \end{aligned}$$

Asociativnost (iii) produkta dokažemo s skokom v predikatni račun.

$$\begin{aligned} a(R*S)*Tb &\sim \exists x\,(a(R*S)x \text{ in } xTb) \\ &\sim \exists x\,(\exists y\,(aRy \text{ in } ySx) \text{ in } xTb) \\ &\sim \exists x\,\exists y\,(aRy \text{ in } ySx \text{ in } xTb) \\ &\sim \exists y\,\exists x\,(aRy \text{ in } ySx \text{ in } xTb) \\ &\sim \exists y\,(aRy \text{ in } \exists x\,(ySx \text{ in } xTb)) \\ &\sim \exists y\,(aRy \text{ in } yS*Tb) \\ &\sim aR*(S*T)b \end{aligned}$$

Torej sta relaciji $(R*S)*T$ in $R*(S*T)$ enaki.

Za utemeljitev (iv) računamo takole:

$$\begin{aligned} aR*(S\cup T)b &\sim \exists x\,(aRx \text{ in } x(S\cup T)b) \\ &\sim \exists x\,(aRx \text{ in } (xSb \text{ ali } xTb)) \\ &\sim \exists x\,((aRx \text{ in } xSb) \text{ ali } (aRx \text{ in } xTb)) \\ &\sim \exists x\,(aRx \text{ in } xSb) \text{ ali } \exists x\,(aRx \text{ in } xTb) \\ &\sim (aR*Sb) \text{ ali } (aR*Tb) \\ &\sim a(R*S)\cup(R*T)b \end{aligned}$$

Distributivnost pri množenju z desne strani uženemo na simetričen način.

Lastnosti (v) in (vi) lahko pokažemo z rutinskima računoma, zato dokaza izpustimo. □

Na tem mestu naj omenimo, da distributivnost produkta preko preseka relacij ne velja. V splošnem relaciji $R*(S\cap T)$ in $(R*S)\cap(R*T)$ nista enaki. Najenostavnejši protiprimer je predstavljen na sliki 4.3, kjer so na istem grafu predstavljene relacije R, S in T.

Zaradi asociativnosti množenja relacij smemo definirati potence relacij. Naj bo R relacija v A. Potence relacije R z nenegativnimi eksponenti definiramo z rekurzivnim opisom.

$$\begin{aligned} R^0 &= \mathrm{id}_A, \\ R^{n+1} &= R*R^n, \quad \text{če je } n \geq 0. \end{aligned}$$

Izračunajmo nekaj zaporednih potenc relacije R.

$$\begin{aligned} R^1 &= R*R^0 = R*\mathrm{id}_A = R, \\ R^2 &= R*R^1 = R*R, \\ R^3 &= R*R^2 = R*(R*R) = R*R*R, \end{aligned}$$

pri čemer zadnja enakost velja zaradi asociativnosti množenja relacij.

Potence relacije z negativnimi eksponenti definiramo kot potence njej inverzne relacije s pozitivnimi eksponenti. Za $n > 0$ naj bo

$$R^{-n} = (R^{-1})^n.$$

Pri tem velja

$$R^m * R^n = R^{m+n}, \quad \text{če sta števili } m \text{ in } n \text{ istega znaka.}$$

Nasprotno predznačenih eksponentov ne smemo seštevati. Najenostavnejši zgled konstruiramo v družini sorodstvenih relacij med ljudmi, relaciji "*otrok*" in "*roditelj*" sta si paroma inverzni, saj je

$$a \textit{ otrok } b \quad \text{natanko tedaj, ko je} \quad b \textit{ roditelj } a.$$

Pišemo lahko

$$\textit{otrok}^{-1} = \textit{roditelj}.$$

V množici ljudi izberimo sestri c in d, ki imata skupnega očeta. Zato velja

$$\exists x(c \textit{ otrok } x \text{ in } x \textit{ roditelj } d)$$

in odtod

$$c \textit{ otrok}*\textit{roditelj } d \;\sim\; c \textit{ otrok}^1 * \textit{otrok}^{-1} \; d.$$

Ker sta c in d različna človeka, relacija $\textit{otrok}^1 * \textit{otrok}^{-1}$ ni enaka relaciji identitete $\textit{otrok}^0$ v množici ljudi.

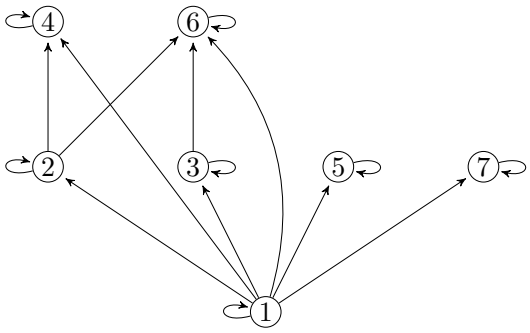
4.3 Grafična predstavitev relacije in potence

Naj bo R relacija v *končni* množici A. Relacijo R lahko predstavimo grafično, elemente množice A predstavimo kot točke/krožce v ravnini. Za vsak par a, b, za katerega je aRb, narišemo usmerjeno povezavo (puščico) od točke, ki predstavlja a, do točke, ki predstavlja b.

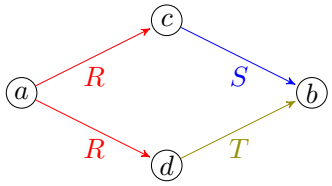
Relacijo $R = \{(a, b), (b, c), (c, d), (c, c)\}$ v množici $A = \{a, b, c, d\}$ predstavimo z grafom na sliki 4.1. Par $(c, c) \in R$ predstavimo z *zanko*, usmerjeno povezavo z začetkom in koncem v isti točki.

Slika 4.1: Graf relacije $R = \{(a, b), (b, c), (c, d), (c, c)\}$.

Na sliki 4.2 je prikazan graf relacije deljivosti | na množici števil $\{1, \ldots, 7\}$.



Slika 4.2: Graf relacije deljivosti na $\{1, \ldots, 7\}$.



Slika 4.3: Velja $a((R * S) \cap (R * T))b$, toda $\neg(aR * (S \cap T)b)$.

Z grafom relacije relativno enostavno konstruiramo zgled, zakaj produkt relacij ne distribuira preko preseka relacij. Lastnost trditve 4.1(iv) za presek ne velja, glej sliko 4.3.

Kako lahko iz grafa relacije R odločimo, katere izmed lastnosti ima relacija R? Ali lahko lastnost relacije opišemo tudi algebraično? Refleksivnost na grafu preverimo enostavno: relacija je refleksivna natanko tedaj, ko ima njen graf v vsaki točki zanko. Oziroma, relacija R v A je refleksivna natanko tedaj, ko vsebuje relacijo id_A.

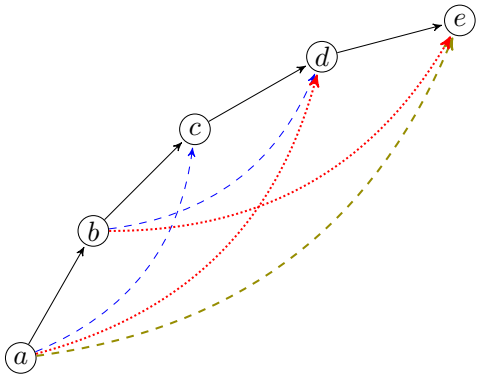
Denimo, da je R simetrična relacija v A in $a, b \in A$. Potem velja bodisi aRb in bRa ali pa $\neg aRb$ in $\neg bRa$. V grafu simetrične relacije R sta med vsakim parom točk prisotni ali povezavi v obeh smereh ali pa med njima ni povezave. Simetrija relacije ni občutljiva na zanke, te so lahko prisotne ali pač ne. Enakovredno, relacija R je simetrična natanko tedaj, ko sovpada s svojo inverzno relacijo R^{-1}.

Par nasprotno usmerjenih povezav v grafu relacije ponavadi nadomestimo z eno neusmerjeno povezavo, glej sliko 4.4.

Antisimetričnost relacije R je iz grafa razvidna z odsotnostjo paroma nasprotno usmerjenih povezav. Če lahko najdemo par različnih točk $a, b \in A$, za kateri v grafu relacije R obstaja par nasprotno usmerjenih povezav s krajiščema v a in b, potem velja aRb, bRa in tudi $a \neq b$. Takšna relacija ni antisimetrična.

Par nasprotno usmerjenih povezav relacije R pripada preseku $R \cap R^{-1}$. Antisimetrija

Slika 4.4: Zvezo aRb in bRa lahko predstavimo z neusmerjeno povezavo.



Slika 4.5: Zgled potence relacije R, prikazani so $aR^2c, bR^2d, aR^3d, bR^3e$ in aR^4e.

relacije R zagotavlja, da je $R \cap R^{-1} \subseteq \mathrm{id}_A$.

Kot v primeru simetrije tudi antisimetrija ni občutljiva na prisotnost oz. odsotnost zank v grafu relacije R.

Relacija R je sovisna natanko tedaj, ko med poljubnima različnima elementoma $a, b \in A$ v grafu relacije R obstaja vsaj ena povezava. To je enakovredno dejstvu, da vsak urejen par (a, b) pripada relaciji R ali pripada relaciji R^{-1} ali pa sta njegovi koordinati enaki. Enakovredno, $\mathrm{id}_A \cup R \cup R^{-1} = U_A$.

Relacija R je enolična, če iz vsake točke a v grafu relacije R izhaja *največ* ena puščica. Denimo, da velja aRb in aRc, pri čemer sta b in c različna. Torej je $cR^{-1}a$ in aRb in odtod lahko sklepamo, da par *različnih* elementov množice A pripada relaciji $R^{-1} * R$. Če je relacija R enolična, se to ne more zgoditi. Za enolične relacije torej velja $R^{-1} * R \subseteq \mathrm{id}_A$.

Tranzitivnost relacije R zahteva nekoliko drugačen pristop. Če je aRz in zRb, po definiciji produkta relacij velja $aR * Rb$ oziroma aR^2b. Če je relacija R tranzitivna, potem za vsak par a, b iz aR^2b sledi aRb. To je enakovredno vsebovanosti

$$R^2 \subseteq R.$$

Relacijo R^2 v grafu predstavimo z uporabo dveh zaporednih puščic relacije R. Če se lahko iz točke a z uporabo dveh zaporednih puščic relacije R sprehodimo do b, potem je aR^2b. V tranzitivni relaciji imamo torej za vsak par zaporednih puščic tudi direktno

bližnjico.

Skočimo za trenutek na sliko 4.5. Če je relacija R tranzitivna ter velja aRb in bRc, potem mora po tranzitivnosti veljati tudi aRc. Iz aRc in cRd po tranzitivnosti sledi tudi aRd, in tako naprej. Lahko bi rekli, da v grafu tranzitivne relacije ne smemo izpustiti nobene od *bližnjic*.

Spodnjo trditev lahko zapišemo brez dokaza.

Trditev 4.2 *Naj bo R relacija v množici A. Za $a, b \in A$ in poljuben $n \in \mathbb{N}$ je*

$$aR^n b$$

natanko tedaj, ko lahko v grafu relacije R iz a pridemo do b z uporabo n zaporednih puščic.

Za konec razdelka zapišimo še algebraično karakterizacijo lastnosti relacij, ki smo jo implicitno ravnokar izpeljali.

Trditev 4.3 [Algebraična karakterizacija lastnosti relacij]
Relacija R v A je

- *refleksivna natanko tedaj, ko je* $\mathrm{id}_A \subseteq R$
- *simetrična natanko tedaj, ko je* $R = R^{-1}$,
- *antisimetrična natanko tedaj, ko je* $R \cap R^{-1} \subseteq \mathrm{id}_A$,
- *tranzitivna natanko tedaj, ko je* $R^2 \subseteq R$,
- *sovisna natanko tedaj, ko je* $\mathrm{id}_A \cup R \cup R^{-1} = U_A$,
- *enolična natanko tedaj, ko je* $R^{-1} * R \subseteq \mathrm{id}_A$.

4.4 Ovojnice relacij

Naj bo $\mathcal{L}$ lastnost relacij in R relacija v množici A. Relacija R morda lastnosti $\mathcal{L}$ nima, morda pa bi z dodajanjem parov relaciji R uspeli doseči, da bo razširjena relacija imela lastnost $\mathcal{L}$.

Po vsebovanosti najmanjšo takšno relacijo bomo imenovali *$\mathcal{L}$-ovojnica* relacije R.

Natančneje, relacija $R_{\mathcal{L}}$ v isti množici A je *$\mathcal{L}$-ovojnica* relacije R, če je

(CL1) $R \subseteq R_{\mathcal{L}}$,

(CL2) $R_{\mathcal{L}}$ ima lastnost $\mathcal{L}$ in

(CL3) če za relacijo S velja, da $R \subseteq S$ in ima S lastnost $\mathcal{L}$, potem je $R_{\mathcal{L}} \subseteq S$.

Vsaka lastnost relacij ne dopušča ovojnice. Naj bo R relacija v A in naj obstajata različna $a, b \in A$, za katera velja tako aRb kot bRa. Relacija R ni antisimetrična, elementa a in b predstavljata protiprimer. Ravno tako nobena relacija R', za katero velja $R \subseteq R'$, ni antisimetrična. Recikliramo lahko isti protiprimer.

Ne glede na lastnost $\mathcal{L}$, mora biti $\mathcal{L}$-ovojnica univerzalne relacije U_A zaradi (CL1) enaka sama sebi in po (CL2) mora imeti lastnost $\mathcal{L}$.

Denimo, da ima relacija R samo pet parov, z $\mathcal{L}_6$ pa označimo lastnost "*vsebovati kot element natanko 6 različnih parov.*" Relacijo s 5 elementi lahko na veliko[2] različnih načinov dopolnimo do relacije z natančno 6 elementi. Nobena od tako dobljenih razširjenih relacij pa ni $\mathcal{L}_6$-ovojnica relacije R, saj bi v skladu z (CL3) morala biti vsebovana v preostalih.

Trditev 4.4 *Naj bo $\mathcal{L}$ lastnost relacij in A poljubna množica. Denimo, da $\mathcal{L}$ zadošča naslednjima pogojema.*

(C1) *univerzalna relacija U_A ima lastnost $\mathcal{L}$ in*

(C2) *$\mathcal{L}$ je zaprta za preseke relacij — presek družine relacij, ki vse zadoščajo $\mathcal{L}$, ima tudi lastnost $\mathcal{L}$.*

Potem za vsako relacijo R v množici A obstaja njena $\mathcal{L}$-ovojnica $R_{\mathcal{L}}$.

Dokaz. Označimo z D družino vseh relacij v množici A, ki imajo lastnost $\mathcal{L}$ in vsebujejo relacijo R. Po (C1) je $U_A \in D$, torej je D neprazna družina relacij in obstaja njen presek $P = \bigcap D$. Trdimo, da je P $\mathcal{L}$-ovojnica relacije R. Preverimo pogoje iz definicije $\mathcal{L}$-ovojnice.

Ker vse relacije iz D vsebujejo R, je $R \subseteq P$. Velja torej pogoj (CL1). Iz (C2) sledi, da ima P lastnost $\mathcal{L}$. Torej je izpolnjen tudi pogoj (CL2). Če je $R \subseteq S$ in ima S lastnost $\mathcal{L}$, je $S \in D$ po definiciji družine D. Zato je $P \subseteq S$, z drugimi besedami, tudi pogoj (CL3) je izpolnjen.

Zaradi (CL3) je $\mathcal{L}$-ovojnica relacije R enolična, torej lahko zapišemo $P = R_{\mathcal{L}}$. □

Naj bo R relacija v A. *Refleksivna ovojnica* relacije R je enaka $R \cup \mathrm{id}_A$, njena *simetrična ovojnica* pa je enaka $R \cup R^{-1}$.

Tranzitivno ovojnico relacije R označimo z R^+. Velja

$$R^+ = \bigcup_{k=1}^{\infty} R^k, \tag{4.3}$$

tranzitivna ovojnica relacije R je unija vseh pozitivnih potenc relacije R.

[2] Na vsaj 4 načine. Zakaj?

Preverimo, da je R^+ res tranzitivna ovojnica relacije R. Pogoj (CL1) je izpolnjen, saj je $R = R^1$ prvi člen unije v definiciji R^+.

Denimo, da velja aR^+b in bR^+c. Torej obstajata takšni naravni števili k_1 in k_2, da je $aR^{k_1}b$ in $bR^{k_2}c$. Po definiciji produkta je $aR^{k_1} * R^{k_2}c$ ali enakovredno $aR^{k_1+k_2}c$. Zato je tudi aR^+c in relacija R^+ je tranzitivna. Pod streho je tudi pogoj (CL2).

Denimo, da je $\overline{R}$ relacija, ki zadošča tako (CL1) kot (CL2). Z indukcijo bomo pokazali, da za vsako naravno število $n \geq 1$ velja vsebovanost $R^n \subseteq \overline{R}$. Pogoj (CL1) služi kot baza indukcije in indukcijsko privzemimo, da je $R^{n'} \subseteq \overline{R}$ za vsa naravna števila $1 \leq n' < n$. Računajmo:

$$R^n = R * R^{n-1} \subseteq \overline{R} * \overline{R} \subseteq \overline{R},$$

zadnja vsebovanost sledi zaradi tranzitivnosti relacije $\overline{R}$.

Od tod sklepamo, da je $R^+ \subseteq \overline{R}$ za vsako relacijo $\overline{R}$, ki zadošča (CL1) in (CL2). Zato R^+ ustreza pogoju (CL3).

Če je R tranzitivna, potem mora biti R^+ enaka R. Zaradi tranzitivnosti relacije R velja $R^2 \subseteq R$. Z množenjem vsebovanosti z relacijo R pridelamo tudi $R^3 \subseteq R^2$. Induktivno sledi

$$R \supseteq R^2 \supseteq R^3 \supseteq R^4 \supseteq R^5 \ldots$$

in zato je $R^+ = \cup_{k=1}^{\infty} R^k = R^1 = R$.

Tranzitivno-refleksivna ovojnica relacije R, označimo jo z R^*, je definirana zelo podobno. Velja namreč

$$R^* = \bigcup_{k=0}^{\infty} R^k, \tag{4.4}$$

tranzitivno-refleksivna ovojnica relacije R je unija vseh nenegativnih potenc relacije R.

Utemeljitev je enostavna. Ker je R^* tranzitivna, velja $R^+ \subseteq R^*$. Ker je tudi refleksivna, je $R^0 \subseteq R^*$. Ker je unija $R^+ \cup R^0$ tako tranzitivna kot refleksivna, je enaka tranzitivno-refleksivni ovojnici R^* relacije R.

Končajmo z grafično interpretacijo relacij R^+ in R^*. Kako lahko omenjeni ovojnici *preberemo* iz grafa relacije R?

Zveza aR^kb pomeni, da se lahko v grafu relacije R premaknemo od a do b in pri tem prehodimo natančno k zaporednih puščic (usmerjenih povezav). Zato lahko zapišemo trditev:

Trditev 4.5 *Naj bo R relacija v A in $a, b \in A$. Potem je*

(i) *aR^+b natanko tedaj, ko se lahko v grafu relacije R sprehodimo od a do b in pri tem uporabimo vsaj eno usmerjeno povezavo,*

(ii) *aR^*b natanko tedaj, ko se lahko v grafu relacije R sprehodimo od a do b in pri tem uporabimo 0 ali več usmerjenih povezav.*

4.5 Ekvivalenčna relacija

V tem razdelku bomo spoznali ekvivalenčne relacije. Relacija R v množici A je *ekvivalenčna*, če je

(EQ1) refleksivna,

(EQ2) simetrična in

(EQ3) tranzitivna.

Vsaka od teh treh sestavin sama zase ni posebej pretresljiva. Tranzitivnost smo intuitivno celo povezovali z urejenostmi. Toda uporabljene skupaj v istem receptu bodo skuhale presenetljivo specialiteto[3].

Oglejmo si nekaj zgledov ekvivalenčnih relacij.

- Relacija *vzporednosti* $\parallel$ v družini premic v ravnini. Premici p in q sta vzporedni,

  $$p \parallel q \quad \text{natanko tedaj, ko} \quad \textit{imata } p \textit{ in } q \textit{ isto} \text{ “smer”}.$$

  Tako definirana vzporednost premic je ekvivalenčna relacija. Če vzporednost premic definiramo z disjunktnostjo, imamo po eni strani težave z refleksivnostjo relacije, po drugi strani pa tako definirana vzporednost ne preživi v prostorih višjih dimenzij.

- *Kongruenčna relacija* v množici celih števil. Izberimo *veliko*[4] naravno število $m \geq 1$. Za celi števili a in b pravimo, da sta *kongruentni po modulu* m,

  $$a \equiv b \pmod{m} \quad \text{natanko tedaj, ko } m \textit{ deli} \text{ razliko } a - b.$$

  Kongruenca po modulu m je ekvivalenčna relacija v množici celih števil $\mathbb{Z}$.

- Denimo, da imamo v žepu štiri kovance vrednosti $1, 2, 5$ in 5 evrskih centov. Z njimi lahko natančno plačamo naslednje vrednosti: $1, 2, 3, 5, 6, 7, 8, 10, 11, 12, 13$ in tudi, na trivialen način, 0 centov. Posplošimo na splošna (naraščajoča) zaporedja naravnih števil.

  (Naraščajoči) zaporedji $\alpha = (a_1, \ldots, a_k)$ in $\beta = (b_1, \ldots, b_\ell)$ sta *plačilno enakovredni*, če imata zaporedji α in β isto družino delnih vsot. Tako sta denimo zaporedji $(1, 2, 2)$ in $(1, 1, 1, 1, 1)$ plačilno enakovredni, saj lahko z obema *plačamo* natančno zneske $0, 1, 2, 3, 4$ in 5. Zaporedji $(1, 1, 2)$ in $(1, 3)$ pa nista plačilno enakovredni, saj z drugim zaporedjem ne moremo plačati zneska 2.

  Plačilna enakovrednost je ekvivalenčna relacija v množici (končnih, naraščajočih) zaporedij naravnih števil.

[3] Če nismo pravi navdušenci nad tovrstno kulinariko, bi bilo morda ustrezneje reči za silo užitno kosilo.

[4] 1 je sicer dovolj veliko število, a bolj zanimivo bi si bilo izbrati kaj strogo večjega od 1. Vsekakor je kongruenca po modulu 2 zelo zanimiva in uporabna. Izbira 0 pa je premajhna.

4.5.1 Ekvivalenčni razredi

Naj bo R ekvivalenčna relacija v množici A. *Ekvivalenčni razred* elementa $a \in A$, $R[a]$, definiramo kot

$$R[a] = \{x \mid xRa\}. \tag{4.5}$$

Ekvivalenčni razred elementa $a \in A$ je množica vseh tistih elementov množice A, ki so v relaciji z a.

Zaradi refleksivnosti relacije R za vsak ekvivalenčni razred $R[a]$ velja $a \in R[a]$. Vsak element $a \in A$ pripada *svojemu* ekvivalenčnemu razredu. Posledično to pomeni, da so ekvivalenčni razredi neprazni.

Faktorska ali *kvocientna množica* množice A po ekvivalenčni relaciji R, označimo jo z A/R, je družina vseh ekvivalenčnih razredov,

$$A/R = \{R[a] \mid a \in A\}. \tag{4.6}$$

Na prvi pogled se zdi, da je ekvivalenčnih razredov v faktorski množici natančno toliko kot elementov množice A. To je tipično daleč od resnice, ekvivalenčni razredi različnih elementov večkrat sovpadajo. Ta fenomen znamo celo karakterizirati.

Trditev 4.6 *Naj bo R ekvivalenčna relacija v množici A in $a, b \in A$. Potem je*

$$R[a] = R[b] \quad \textit{natanko tedaj, ko je} \quad aRb.$$

Dokaz. ($\Longrightarrow$) Privzemimo enakost ekvivalenčnih razredov $R[a] = R[b]$. Iz $a \in R[a]$ sledi pripadnost $a \in R[b]$ in po definiciji ekvivalenčnega razreda tudi aRb.

($\Longleftarrow$) Privzemimo, da je aRb. Enakost ekvivalenčnih razredov $R[a] = R[b]$ bomo ugnali tako, da bomo pokazali obe vsebovanosti $R[a] \subseteq R[b]$ in $R[b] \subseteq R[a]$.

Izberimo najprej poljuben element $c \in R[a]$. Po definiciji velja cRa, po predpostavki pa aRb. Ker je relacija R tranzitivna, lahko sklepamo na cRb oziroma na $c \in R[b]$. Torej je $R[a] \subseteq R[b]$.

Sedaj izberimo poljuben element $c \in R[b]$, torej je cRb. Ker je R simetrična, lahko predpostavko enakovredno prepišemo kot bRa, z uporabo tranzitivnosti pa dobimo cRa. Torej je $c \in R[a]$ in zato $R[b] \subseteq R[a]$. Dokaz je zaključen. □

Glavni rezultat razdelka je naslednji izrek.

Izrek 4.7 *Naj bo R ekvivalenčna relacija v množici A. Potem je kvocientna množica A/R* razbitje *množice A.*

Dokaz. Pokazati je potrebno, da je kvocientna množica A/R pokritje množice A, pogoj (R1), da so bloki razbitja neprazni (R2) in paroma disjunktni (R3), glej definicijo razbitja na strani 77.

Izberimo poljuben element $b \in A$. Ker $b \in R[b]$, velja tudi

$$b \in \bigcup_{a \in A} R[a],$$

zato (R1). Ker so ekvivalenčni razredi neprazni, velja tudi pogoj (R2).

Za (R3) je dovolj pokazati implikacijo

$$\text{če je} \quad R[a] \cap R[b] \neq \emptyset, \quad \text{potem je} \quad R[a] = R[b].$$

Denimo, da $c \in R[a] \cap R[b]$. Torej velja cRa in cRb. Z uporabo simetrije in tranzitivnosti lahko izpeljemo aRb, z uporabo trditve 4.6 pa tudi enakost $R[a] = R[b]$ in dokaz je končan. □

Ekvivalenčno relacijo v množici A smemo zaradi izreka 4.7 *enačiti* z razbitjem množice A. Ekvivalenčna relacija R definira razbitje A/R. Po drugi strani pa vsako razbitje množice A porodi ekvivalenčno relacijo[5] R z opisom

$$aRb \quad \text{natanko tedaj, ko } a \text{ in } b \text{ pripadata istemu bloku razbitja.}$$

Za konec si oglejmo še zgled kvocientne množice v primeru kongruenčne relacije v množici celih števil pri $m = 3$. Ekvivalenčni razred elementa a bomo označili kar z $[a]$.

Katera števila so v relaciji s številom $0, 1$ oziroma 2? To so natančno tista števila, ki dajo pri deljenju s 3 ostanek $0, 1$ oziroma 2. Zato lahko zapišemo tudi ekvivalenčne razrede:

$$[0] = \{\ldots, -3, 0, 3, 6, \ldots\}$$
$$[1] = \{\ldots, -2, 1, 4, 7, \ldots\}$$
$$[2] = \{\ldots, -1, 2, 5, 8, \ldots\}$$

in kvocientno množico

$$\mathbb{Z}/\,(\mathrm{mod}\ 3) = \{[0], [1], [2]\}.$$

4.6 Relacije urejenosti

V zadnjem razdelku poglavja se bomo lotili relacij urejenosti. Tranzitivnost je osnovna lastnost, ki jo srečamo pri vseh vrstah urejenosti.

V številskih množicah, če odmislimo kompleksna števila, poznamo urejenost po velikosti. Vse relacije $\leq, \geq, <, >$ pa bomo smeli dojemati kot več plati iste medalje, čeprav je, denimo, relacija $\leq$ refleksivna, njena stroga varianta pa ne.

[5] Da je tako definirana relacija res ekvivalenčna, lahko pokaže bralec sam.

4.6.1 Delna in linearna urejenost

Delna urejenost v množici A je *vsaka* relacija v A, ki je

(PO1) refleksivna,

(PO2) tranzitivna in

(PO3) antisimetrična.

Dogovorimo se še za terminologijo. Delno urejenost v množici A bomo označevali z generično oznako $\preccurlyeq$ in dejali, da $\preccurlyeq$ delno ureja množico A. Poleg tega

- $a \preccurlyeq b$ preberemo z "*a je pod b*",
- $a \prec b$ pomeni isto kot $a \preccurlyeq b$ in $a \neq b$, kar preberemo kot "*a je strogo pod b*",
- $a \succcurlyeq b$ in $a \succ b$ pomenita isto kot $b \preccurlyeq a$ in $b \prec a$, zvezi preberemo tudi kot "*a je nad b*" oziroma "*a je strogo nad b*",
- zapis $a \prec b \preccurlyeq c$ pomeni konjunkcijo $a \prec b$ in $b \preccurlyeq c$.

Navedimo še nekaj zgledov delnih urejenosti:

1. Najbolj tipičen zgled delne urejenosti je relacija vsebovanosti $\subseteq$ v družini množic.
2. Relacija deljivosti $|$, kjer $a|b$ preberemo kot "*a deli b*", je delna urejenost v množici naravnih števil.

   Deljivost *ni* delna urejenost v množici celih števil, saj ni antisimetrična. Števili 5 in -5 sta vzajemno deljivi, nista pa enaki.
3. Če izjavni izraz I nastopa v izjavnem izrazu J, potem to označimo z $I \hookrightarrow J$. Relacija $\hookrightarrow$ je delna urejenost v množici vseh izjavnih izrazov.
4. Relaciji $\leq$ in $\geq$ sta delni urejenosti v množicah $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ oziroma $\mathbb{R}$.

Naj bo $\preccurlyeq$ delna urejenost v množici A. Če za $a, b \in A$ velja $a \preccurlyeq b$ ali $b \preccurlyeq a$, potem za a in b pravimo, da sta *primerljiva*.

Če za delno urejenost $\preccurlyeq$ v množici A velja, da sta vsaka dva elementa iz A primerljiva, potem pravimo, da je $\preccurlyeq$ *linearna urejenost* v A, oziroma da $\preccurlyeq$ *linearno ureja* množico A.

Z drugimi besedami, linearna urejenost je takšna delna urejenost, ki je celo sovisna.

Navedimo dva zgleda:

1. relacija $\leq$ linearno ureja vsako podmnožico realnih števil $\mathbb{R}$,
2. črke slovenske abecede so *linearno urejene po abecedi*.

Zapišimo še rezultat o zožitvi urejenosti.

Trditev 4.8 *Denimo, da $\preccurlyeq$ delno ureja množico A in je $A' \subseteq A$. Potem zožitev relacije $\preccurlyeq$ delno ureja tudi A'.*

Če je $\preccurlyeq$ linearna urejenost v A, potem zožitev celo linearno ureja A'.

Dokaz. Upoštevati je potrebno samo, da je vsaka izmed lastnosti refleksivnost, antisimetričnost, tranzitivnost in tudi sovisnost opisana z univerzalno izjavno formulo. □

Na tem mestu naj omenimo, da lahko delna urejenost množice A celo linearno ureja katero od njenih podmnožic. Takšni podmnožici V pravimo tudi *veriga*. Kot primer navedimo, da relacija deljivosti linearno ureja družino potenc števila 2 v množici naravnih števil $\mathbb{N}$ — $\{1, 2, 4, 8, \ldots, 2^k, \ldots\}$ je veriga za deljivost.

Denimo, da relacija $\preccurlyeq_A$ ureja množico A in relacija $\preccurlyeq_B$ ureja množico B. V kartezičnem produktu $A \times B$ lahko definiramo *leksikografsko urejenost* $\preccurlyeq_{lex}$ z naslednjim opisom:

$$(a, b) \preccurlyeq_{lex} (a', b')$$

natanko tedaj, ko je

(i) $a \prec_A a'$ ali

(ii) $a = a'$ in $b \preccurlyeq_B b'$.

Urejeno n-terico smemo enačiti z zaporedjem dolžine n. Leksikografsko urejenost lahko posplošimo tudi na končna zaporedja morda celo različnih dolžin.

Denimo, da sta $\alpha = (a_1, a_2, \ldots, a_m)$ in $\beta = (b_1, b_2, \ldots, b_n)$ različni zaporedji. Potem velja

$$\alpha \prec_{lex} \beta$$

natanko tedaj, ko bodisi za najmanjši indeks i, pri katerem je $a_i \neq b_i$, velja $a_i \prec b_i$ bodisi se zaporedji α in β ujemata v prvih m členih in je $n > m$.

Navedimo poenostavljeno verzijo izreka o leksikografski urejenosti.

Izrek 4.9 *Naj bo množica A delno urejena z relacijo $\preccurlyeq$. Potem leksikografska urejenost $\preccurlyeq_{lex}$ delno ureja družino vseh končnih zaporedij elementov iz A.*

Če je $\preccurlyeq$ linearna urejenost v A, potem $\preccurlyeq_{lex}$ celo linearno ureja družino vseh končnih zaporedij iz A.

Izrek 4.9 seveda deluje tudi na urejenih parih, ki jih lahko interpretiramo kot zaporedja dolžine 2.

Linearna urejenost črk "*po abecedi*" porodi običajno leksikografsko (tudi *slovarsko*) urejenost besed v slovarju, ki jih "po abecedi" primerjamo na prvem mestu, na katerem se besede razlikujejo.

4.6.2 Posebni elementi

V tem razdelku definirajmo nekatere posebne elemente v delnih urejenostih. Naj bo $\preccurlyeq$ delna urejenost v A. Za $a \in A$ velja

(i) a je *minimalni element*, če za vsak $a' \in A$ velja $a' \not\prec a$,

(ii) a je *prvi element*, če za vsak $a' \in A$ velja $a \preccurlyeq a'$,

(iii) a je *maksimalni element*, če za vsak $a' \in A$ velja $a \not\prec a'$ in

(iv) a je *zadnji element*, če za vsak $a' \in A$ velja $a' \preccurlyeq a$.

Maksimalni in zadnji elementi so natančno minimalni in prvi elementi v *inverzni urejenosti*, tj. urejenosti z inverzno relacijo. Naslednji lastnosti bomo zapisali zgolj za prve oz. minimalne elemente, smiselno pa veljajo tudi za zadnje oz. maksimalne.

Naj bo $\preccurlyeq$ delna urejenost v A. Veljata naslednji lastnosti.

- Če je a prvi element v A, potem je a tudi edini minimalni element v A. Posebej, v delni urejenosti obstaja kvečjemu en prvi element.
- Če je a minimalni element v *linearni urejenosti* $\preccurlyeq$, potem je a tudi prvi element.

Delna urejenost dopušča več minimalnih elementov. V družini *nepraznih podmnožic* množice A glede na vsebovanost so vsi *singletoni* minimalni elementi.

V splošnem delna urejenost shaja celo brez minimalnih ali prvih elementov. Množica realnih števil nima minimalnega elementa glede na standardno urejenost $\leq$.

Naj bo $\preccurlyeq$ delna urejenost v množici A in $A' \subseteq A$. Pravimo, da je A' *navzgor omejena*, če obstaja tak $\overline{a} \in A$, da za vsak element $a' \in A'$ velja $a' \preccurlyeq \overline{a}$. Elementu $\overline{a}$ pravimo tudi *zgornja meja* množice A' v A.

Denimo, da je množica A' navzgor omejena in naj bo $\overline{A}$ množica vseh zgornjih mej množice A'. Če obstaja prvi element množice $\overline{A}$ (glede na zožitev urejenosti na $\overline{A}$), ga imenujemo *natančna zgornja meja* (tudi *supremum*) množice A' v A in ga označimo s $\sup A'$. Če velja celo $\sup A' \in A'$, potem mu pravimo tudi *maksimum* množice A' in ga označimo z $\max A'$.

Zgodbo ponovimo tudi na spodnjem robu. Za množico $A^* \subseteq A$ pravimo, da je *navzdol omejena*, če obstaja tak $\underline{a} \in A$, da za vsak element $a^* \in A^*$ velja $\underline{a} \preccurlyeq a^*$. Elementu $\underline{a}$ pravimo tudi *spodnja meja* množice A^* v A.

Če je množica A^* navzdol omejena, potem z $\underline{A}$ označimo množico vseh spodnjih mej množice A^*. Če obstaja zadnji element množice $\underline{A}$ (glede na zožitev urejenosti na $\underline{A}$), ga imenujemo *natančna spodnja meja* (tudi *infimum*) množice A^* v A in ga označimo z $\inf A^*$. Če velja celo $\inf A^* \in A^*$, potem mu pravimo tudi *minimum* množice A^* in ga označimo z $\min A^*$.

4.6.3 Hassejev diagram

Naj bo $\preccurlyeq$ delna urejenost v množici A. Ker je $\preccurlyeq$ tako refleksivna kot tranzitivna, je njena tranzitivno-refleksivna ovojnica $\preccurlyeq^*$ kar enaka originalni relaciji $\preccurlyeq$.

Ali lahko poiščemo *majhno* (v smislu vsebovanosti) relacijo R, za katero sta relaciji R^* in $\preccurlyeq$ enaki?
Ali lahko urejenost predstavimo grafično z bolj preglednim grafom? Odgovor skušamo najti v naslednjih vrsticah.

Naj relacija $\preccurlyeq$ delno ureja množico A. V isti množici A definirajmo relacijo *neposrednega naslednika* oziroma *neposrednega predhodnika* z oznako $\prec\!\cdot$ in naslednjim opisom

$$a \prec\!\cdot\, b \quad \text{natanko tedaj, ko je} \quad a \prec b \ \text{ in } \ \neg\exists x(a \prec x \prec b).$$

Z besedami, a je *neposredni predhodnik* b-ja oziroma b je *neposredni naslednik* a-ja natanko tedaj, ko je a strogo pod b in ne obstaja element $x \in A$, različen od a in b, ki bi bil v urejenosti strogo med a in b.

Oglejmo si nekaj zgledov.

- Glede na standardno številsko urejenost $\leq$ v množici $\mathbb{Z}$ (in tudi $\mathbb{N}$) za števili m in n velja $m \lessdot n$ natanko tedaj, ko je $n = m + 1$.
- V množicah $\mathbb{Q}$ in $\mathbb{R}$ števila nimajo neposrednih naslednikov glede na standardno urejenost $\leq$. Za vsaki dve realni (ali racionalni) števili a, b, za kateri je $a < b$, velja, da aritmetična sredina $\frac{a+b}{2}$ leži strogo vmes.
- Opazujmo relacijo vsebovanosti $\subseteq$ v družini $\mathcal{P}A$. Neposredni nasledniki množice $A' \subseteq A$ so natanko tiste podmnožice množice A, ki jih iz A' dobimo z dodajanjem natanko enega dodatnega elementa. Tako sta množici $\{1, 2, 4\}$ in $\{1, 2, 333\}$ neposredna naslednika množice $\{1, 2\}$.
- Neposredni nasledniki naravnega števila $n \neq 0$ glede na relacijo deljivosti so natanko števila oblike np, kjer je p praštevilo.

Trditev 4.10 *Naj bo $\preccurlyeq$ delna urejenost v končni množici A in $\prec\!\cdot$ relacija neposrednega naslednika glede na $\preccurlyeq$. Potem velja*

(i) *relacija $\prec\!\cdot^*$ je enaka relaciji $\preccurlyeq$ in*

(ii) *če je R prava podmnožica relacije $\prec\!\cdot$, potem relacija R^* ni enaka relaciji $\preccurlyeq$.*

Dokaz. Ker je relacija $\prec\!\cdot$ vsebovana v relaciji $\preccurlyeq$, je tudi njena tranzitivno-refleksivna ovojnica $\prec\!\cdot^*$ vsebovana v $\preccurlyeq$.

Denimo, da za različna elementa a, a' velja $a \prec a'$. Naj bo $V = \{a = a_0, a_1, \ldots, a_n = a'\}$ maksimalna veriga z minimalnim elementom a in maksimalnim elementom a', z indeksi

izbranimi tako, da velja

$$a = a_0 \prec a_1 \prec a_2 \prec \ldots \prec a_n = a'$$

Zaradi maksimalnosti verige V za vsak $i \in \{0, \ldots, n-1\}$ velja $a_i \prec\cdot\, a_{i+1}$. Zato velja $a \prec\cdot^n a'$ in posledično je isti par v tranzitivno-refleksivni ovojnici relacije $\prec\cdot$.

Za dokaz (ii) je dovolj opazovati relacijo R, podmnožico relacije $\prec\cdot$, ki se od relacije $\prec\cdot$ razlikuje v natančno enem paru. Naj bo a, a' edini par različnih elementov množice A, za katerega velja

$$a \prec\cdot\, a' \quad \text{in} \quad \neg aRa'.$$

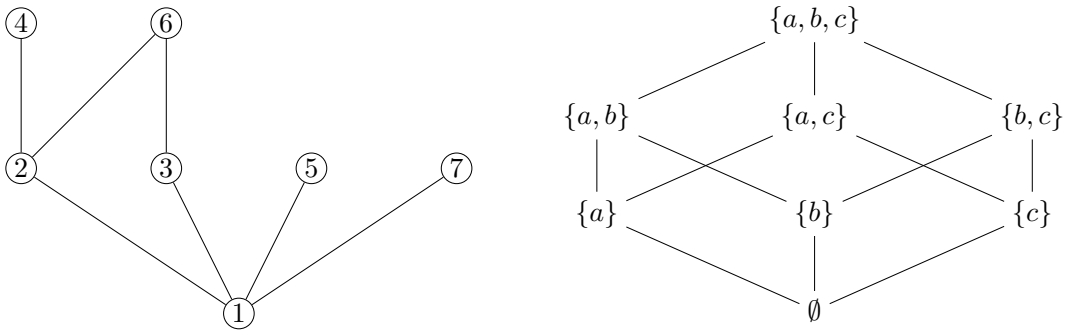
Privzemimo, da je R^* enaka relaciji $\preccurlyeq$. Iz zveze $a \preccurlyeq a'$ sledi, da je aR^*a'. Po definiciji tranzitivno-refleksivne ovojnice obstaja naravno število $n > 0$, za katerega je aR^na'. Velja celo $n \geq 2$, saj $\neg aRa'$.

Ker se relaciji R in $\prec\cdot$ razlikujeta samo v enem urejenem paru, velja tudi $a \prec\cdot^n a'$. To je v protislovju z dejstvom, da je $a \prec\cdot\, a'$, zato R^* ni enaka relaciji $\preccurlyeq$. □

Trditev 4.10 ne velja v primeru urejenosti v neskončnih množicah. Za standardno relacijo urejenosti $\leq$ v množici $\mathbb{Q}$ je relacija neposrednega naslednika prazna relacija.

Hassejev diagram je slikovni prikaz delne urejenosti $\preccurlyeq$ v *končni* množici A. Elemente množice A narišemo kot točke/krožce v ravnini. Za vsak par a, a', za katerega je $a \prec\cdot\, a'$, narišemo naraščajočo povezavo med točkama a in a', pri čemer a narišemo nižje kot a'. Za $a, b \in A$ velja $a \preccurlyeq b$ natanko tedaj, ko lahko v diagramu pridemo od a do b po *vzpenjajoči se poti*.

Na sliki 4.6 sta prikazana Hassejeva diagrama deljivosti v množici $\{1, \ldots, 7\}$ in vsebovanosti v $\mathcal{P}\{a, b, c\}$.



Slika 4.6: Hassejev diagram deljivosti v $\{1, \ldots, 7\}$ in Hassejev diagram vsebovanosti v $\mathcal{P}\{a, b, c\}$.

4.6.4 Dobra urejenost in dobra osnovanost

Poglavje končajmo s še dvema zgledoma relacij urejenosti. Naj bo $\preccurlyeq$ delna urejenost v množici A. Relaciji $\preccurlyeq$ pravimo *dobra osnovanost*, če ima vsaka neprazna podmnožica $A' \subseteq A$ minimalni element.

Relacija $\preccurlyeq$ je *dobra urejenost* množici A, če je $\preccurlyeq$ linearna urejenost in ima vsaka neprazna množica $A' \subseteq A$ minimalni element.

Neskončna padajoča veriga v delno urejeni množici A je neskončno zaporedje

$$a_0, a_1, a_2, a_3, \ldots$$

za katerega za vsak $i \in \mathbb{N}$ velja $a_i \succ a_{i+1}$. Dobro osnovanost lahko karakteriziramo z odsotnostjo neskončnih padajočih verig.

Trditev 4.11 *Za delno urejeno množico A velja, da ima vsaka njena neprazna podmnožica $A' \subseteq A$ minimalni element natanko tedaj, ko A ne vsebuje neskončnih padajočih verig.*

Dokaz. ($\Longrightarrow$) Množica členov neskončne padajoče verige $\{a_0, a_1, \ldots\}$ ne vsebuje minimalnega elementa.

($\Longleftarrow$) Denimo, da je $A_0 \subseteq A$ neprazna množica brez minimalnega elementa. Induktivno bomo konstruirali zaporedje nepraznih množic

$$A_0 \supset A_1 \supset A_2 \supset \ldots,$$

pri čemer nobena od množic A_i, $i \in \mathbb{N}$, ne vsebuje minimalnega elementa, in zaporedje elementov $(a_i)_{i\in\mathbb{N}}, a_i \in A_i$, ki sestavljajo neskončno padajočo verigo

$$a_0 \succ a_1 \succ a_2 \succ \ldots.$$

Element $a_0 \in A_0$ izberimo poljubno in indukcijsko privzemimo, da smo za vsak $i \leq n$ ustrezno določili množico A_i in element a_i. Po predpostavki a_n *ni* minimalni element množice A_n, zato je množica

$$A_{n+1} := \{a \in A_n \mid a \prec a_n\}$$

neprazna. Če je a^* minimalni element množice A_{n+1}, je a^* tudi minimalni element množice A_n. Iz $a \in A_n \setminus A_{n+1}$ in $a \prec a^*$ zaradi $a^* \prec a_n$ namreč po tranzitivnosti relacije $\prec$ sledi $a \in A_{n+1}$, kar je neumnost. Ker A_n po indukcijski predpostavki nima minimalnega elementa, je torej tudi A_{n+1} brez minimalnega elementa. Element $a_{n+1} \in A_{n+1}$ lahko izberemo poljubno. □

Za konec navedimo še nekaj zgledov dobrih urejenosti in dobrih osnovanosti.

1. Relacija $\leq$ dobro ureja množico $\mathbb{N}$ in tudi vsako podmnožico množice $\mathbb{N}$. Leksikografska urejenost pa dobro ureja tudi vsa končna zaporedja naravnih števil.

2. Relacija $\leq$ ni dobra urejenost v množici celih števil. Množica negativnih števil sestavlja neskončno padajočo verigo. Cela števila lahko *dobro uredimo* denimo z uporabo naslednjega naraščajočega zaporedja (ki se sploh ne sklada s standardno urejenostjo celih števil po velikosti):

$$-1, -2, -3, \ldots, 0, 1, 2, 3, \ldots \quad \text{ali pa} \quad 0, -1, 1, -2, 2, \ldots$$

3. Relacija deljivosti je dobra osnovanost v množici vseh naravnih števil. Če sta a, a' različni neničelni naravni števili, za kateri velja $a|a'$, je število a celo strogo manjše od a'. Neskončna padajoča veriga (neničelnih) naravnih števil glede na relacijo deljivosti bi bila tudi neskončna padajoča veriga glede na standardno urejenost naravnih števil po velikosti.

   Zakaj se je potrebno izogibati številu 0? Naravno število 0, ki je deljivo z vsemi naravnimi števili, je celo zadnji element v relaciji deljivosti v $\mathbb{N}$.

4. Tudi nastopanje izjavnih izrazov je dobra osnovanost. Če za različna izraza I, J velja $I \hookrightarrow J$, potem je globina izraza I strogo manjša od globine izraza J. Ker je globina izjavnega izraza naravno število, v urejenosti nastopanja izjavnih izrazov ne obstajajo neskončne padajoče verige.

5. Rekurzivno konstrukcijo izjavnih izrazov, kot smo jo spoznali v prvem poglavju, lahko obravnavamo celo bolj splošno. Rekurzivna konstrukcija kombinatoričnih objektov porodi relacijo R na naraven način. Če za objekt α velja

$$\alpha = \Phi(\alpha_1, \alpha_2, \ldots, \alpha_k),$$

   kar preberemo kot: α konstruiramo s pravilom/postopkom Φ in uporabo predhodno konstruiranih objektov $\alpha_1, \ldots, \alpha_k$, potem naj za vsak $i \in \{1, \ldots, k\}$ velja $\alpha R \alpha_i$. Pri tem želimo, da je tranzitivno-refleksivna ovojnica R^*

   - *antisimetrična*, saj ne želimo cikličnih odvisnosti parov objektov (težava je, če moramo za konstrukcijo objekta α prej poznati β in hkrati za konstrukcijo β poznati α), in
   - *brez neskončnih padajočih verig*, v nasprotnem primeru konstrukcije kombinatoričnega objekta ne moremo rekurzivno prevesti na primitivne/začetne objekte.

   Relacija R^* je torej dobra osnovanost, če naj bo rekurzivni opis kombinatorične konstrukcije ustrezen.

Poglavje 5

Preslikave

V matematični terminologiji se pojma funkcija in preslikava uporabljata včasih celo kot sinonima, ponavadi pa ločeno — nekatere matematične discipline uporabljajo pretežno pojem *funkcija*, druge se raje odločijo za *preslikave*. Pri diskretnih strukturah se bomo odločili za preslikave, pa čeprav bomo govorili tudi o funkcijskih vrednostih preslikave in ne le o njenih slikah.

Matematična analiza velikokrat govori o realnih funkcijah realne spremenljivke — predpisih, ki kot vhodni podatek zgrabijo realno število in kot rezultat ravno tako vrnejo realno število. Logaritemska funkcija je takšen primer — ima pa težave s prebavo negativnih argumentov in tudi ničle. Logaritmiramo lahko le pozitivna realna števila.

V naši obravnavi preslikav se bomo takšnim težavam izognili tako, da bomo možne vhodne podatke preslikave omejili vnaprej. Pojem preslikave bo vseboval podatek o njenem definicijskem območju.

5.1 Preslikave in njihove lastnosti

Relacija $f \subseteq A \times B$ je *preslikava iz A v B*, če je

(P1) f enolična in

(P2) $\mathcal{D}_f = A$, definicijsko območje f je cela množica A.

Če je f preslikava iz A v B, potem pišemo

$$f : A \to B. \tag{5.1}$$

Če je f preslikava (ali celo samo enolična relacija), potem bomo namesto relacijskega zapisa afb uporabljali funkcijski zapis

$$b = f(a). \tag{5.2}$$

V tem primeru tudi pravimo, da f *preslika* element a v element b in tudi, da je b *slika* ali *funkcijska vrednost* a-ja s preslikavo f.

Ker je f enolična (kot relacija), pri poljubnem $a \in \mathcal{D}_f$ obstaja natančno en element b, za katerega je $(a, b) \in f$, oziroma afb. Zato je zapis (5.2) smiseln.

Družino vseh preslikav iz A v B označimo z B^A,

$$B^A = \{f \mid f : A \to B\}. \tag{5.3}$$

Naj bo $f : A \to B$. Pravimo, da je

- f *injektivna* natanko tedaj, ko velja

  $$\forall x \forall y \, (f(x) = f(y) \Rightarrow x = y),$$

  pri tem enakost $f(x) = f(y)$ pomeni isto kot $\exists u \, (u = f(x) \text{ in } u = f(y))$,
- f *surjektivna* natanko tedaj, ko je

  $$\mathcal{Z}_f = B,$$

- f *bijektivna* natanko tedaj, ko je injektivna in surjektivna hkrati.

Injektivnost preslikave f včasih opišemo tudi z enakovredno formulo

$$\forall x \forall y \, (x \neq y \Rightarrow f(x) \neq f(y))$$

lahko pa celo z uporabo ekvivalence

$$\forall x \forall y \, (f(x) = f(y) \Leftrightarrow x = y)$$

saj je za vse pare x, y iz definicijskega območja preslikave f resnična implikacija

$$x = y \Rightarrow f(x) = f(y).$$

Injektivno, surjektivno oziroma bijektivno preslikavo bomo imenovali tudi *injekcija*, *surjekcija* oziroma *bijekcija*.

Naj bo $A = \{a, b, c, d, e\}$, $B = \{1, 2, 3, 4\}$ in $C = \{p, q, r, s, t\}$. Preslikavi $f : A \to B$ in $g : B \to C$ sta formalno definirani takole:

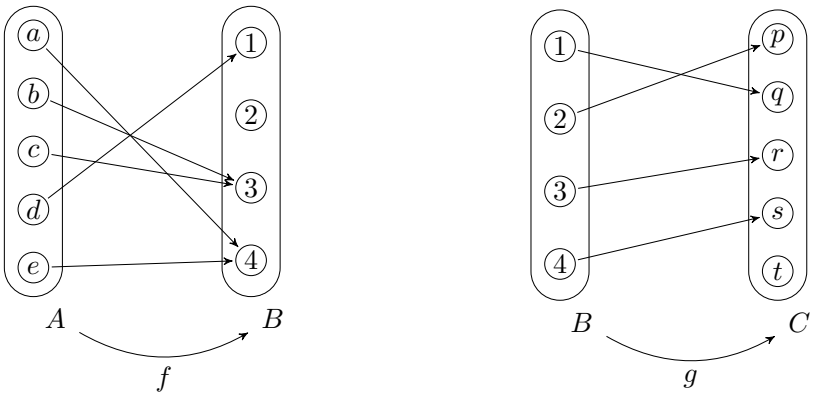
$$f = \{(a, 4), (b, 3), (c, 3), (d, 1), (e, 4)\} \quad \text{in} \quad g = \{(1, q), (2, p), (3, r), (4, s)\}. \tag{5.4}$$

Raje kot z naštevanjem urejenih parov preslikavo predstavimo s tabelaričnim opisom

x	a	b	c	d	e
$f(x)$	4	3	3	1	4

x	1	2	3	4
$g(x)$	q	p	r	s

(5.5)



Slika 5.1: Prikaza preslikav $f : A \to B$ in $g : B \to C$.

ali pa celo prikažemo grafično kot na sliki 5.1.

Preslikava f ni injektivna, saj je $f(b) = f(c) = 3$. Ravno tako f ni surjektivna, število 2 namreč ne pripada njeni zalogi vrednosti $\mathcal{Z}_f = \{1, 3, 4\}$.

Po drugi strani je g injektivna, saj se nobena dva elementa iz množice B ne preslikata v isto sliko v C. Intuitivno bi rekli, da se slike elementov v spodnji vrstici opisa preslikave g ne ponavljajo. Znova pa g ni surjektivna, saj t ne pripada njeni zalogi vrednosti $\mathcal{Z}_g$.

Oglejmo si še nekaj zgledov preslikav:

1. Naj bo A poljubna neprazna množica. Preslikava

   $$\mathrm{id}_A : A \to A$$

   definirana z opisom $\mathrm{id}_A(a) = a$, za vsak $a \in A$, je bijektivna preslikava. Imenujemo jo *preslikava identitete na A*.

2. Naj bo $h : A \to C$ in $A' \subseteq A$. *Zožitev preslikave* h na A' je preslikava

   $$h|_{A'} = h \cap (A' \times C).$$

   Njeno definicijsko območje je enako množici A', za vsak $a \in A'$ pa velja $h|_{A'}(a) = h(a)$. Glej tudi sliko 5.2.

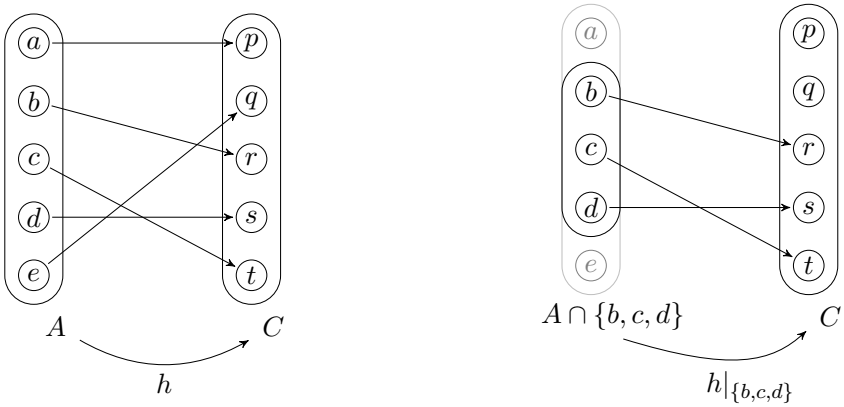
3. Naj bo $A_1 \times A_2 \times \cdots \times A_n$ neprazen kartezični produkt. *Projekcija na i-to komponento* je preslikava

   $$p_i : A_1 \times A_2 \times \cdots \times A_n \to A_i$$

   definirana z opisom

   $$p_i((a_1, a_2, \ldots, a_n)) = a_i.$$

   Projekcija p_i je surjektivna preslikava.



Slika 5.2: Preslikava $h : A \to C$ in njena zožitev $h|_{\{b,c,d\}} : \{b, c, d\} \to C$.

4. Izberimo $B' \subseteq B$. *Vložitev* množice B' v B je preslikava $i : B' \to B$, definirana kot zožitev preslikave identitete,

$$i = \mathrm{id}_B|_{B'}.$$

Vložitev $i : B' \to B$ je injektivna preslikava.

5. Naj bo R ekvivalenčna relacija na množici A. *Naravna projekcija* je preslikava

$$p : A \to A/R$$

definirana z opisom $p(a) = R[a]$, slika posameznega elementa je kar njegov ekvivalenčni razred. Naravna projekcija je surjektivna preslikava.

5.2 Inverzna preslikava in kompozitum

Preslikava $f : A \to B$ je poseben primer relacije. Za vsako relacijo f lahko konstruiramo njej inverzno relacijo f^{-1}. Ali je relacija f^{-1} tudi preslikava?

Odgovor ne bo vedno pozitiven. Inverzni relaciji f^{-1} in g^{-1} preslikav f in g, definiranih v (5.5), sta prikazani na sliki 5.3. Relacija f^{-1} ni enolična, saj velja $3f^{-1}b$ in $3f^{-1}c$. Relacija g^{-1} je enolična, toda ni preslikava iz C v B, saj $t \notin \mathcal{D}_{g^{-1}}$.

Ovira za enoličnost relacije f^{-1} je par puščic iz f, ki se končata v isti točki. To pa je hkrati tudi ovira za injektivnost preslikave f. Po drugi strani pa g^{-1} ni preslikava iz C v B zaradi tega, ker zaloga vrednosti preslikave $\mathcal{Z}_g$ ni enaka celotni množici C.

Tudi v splošnem gre za ista tipa ovir.

Izrek 5.1 *Naj bo $f : A \to B$.*

(i) *Relacija f^{-1} je enolična natanko tedaj, ko je f injektivna.*

(ii) *f^{-1} je preslikava iz B v A natanko tedaj, ko je f bijektivna.*

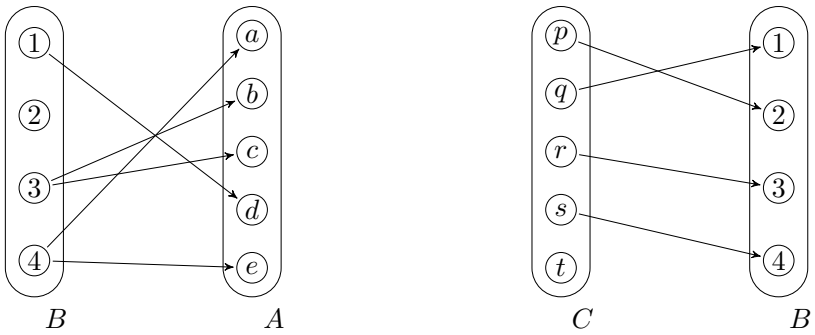
Dokaz. Preslikava f je podmnožica kartezičnega produkta $A \times B$. Zato je $f^{-1} \subseteq B \times A$. Za dokaz (i) je treba premisliti, da je f^{-1} funkcija (enolična relacija) natanko tedaj, ko je f injektivna. Računajmo:

$$\begin{aligned} f^{-1} \text{ je enolična} &\sim \forall x \forall y \forall z\, (z f^{-1} x \text{ in } z f^{-1} y \;\Rightarrow\; x = y) \\ &\sim \forall x \forall y \forall z\, (x f z \text{ in } y f z \;\Rightarrow\; x = y) \\ &\sim \forall x \forall y \forall z\, (z = f(x) \text{ in } z = f(y) \;\Rightarrow\; x = y) \\ &\sim \forall x \forall y\, (f(x) = f(y) \;\Rightarrow\; x = y) \\ &\sim f \text{ je injektivna} \end{aligned}$$

Sledi dokaz točke (ii). Ob tem privzamemo veljavnost (i).

$$\begin{aligned} f^{-1} : B \to A &\sim f^{-1} \text{ je enolična in } \mathcal{D}_{f^{-1}} = B \\ &\sim f \text{ je injektivna in } \mathcal{Z}_f = B \\ &\sim f \text{ je injektivna in } f \text{ je surjektivna} \\ &\sim f \text{ je bijektivna} \end{aligned}$$

Bistvo dokaza točke (ii) leži v opazki, da se definicijsko območje f^{-1} ujema z zalogo vrednosti preslikave f. □



Slika 5.3: Inverzni relaciji f^{-1} in g^{-1} — f^{-1} ni funkcija, g^{-1} ni preslikava iz C v B.

Naj bosta f in g preslikavi. *Kompozitum preslikav* $g \circ f$ definiramo kot

$$g \circ f = f * g. \tag{5.6}$$

Kompozitum preslikav sovpada z relacijskim produktom, v katerem zamenjamo vrstni red faktorjev.

Prepričajmo se najprej, da je kompozitum smiselno definiran — da je dobljeni objekt $f \circ g$ preslikava.

Trditev 5.2

(i) *Če sta f in g enolični relaciji, potem je tudi $g \circ f$ enolična in velja*

$$(g \circ f)(a) = g(f(a)).$$

(ii) *Če je $f : A \to B$ in je $g : B \to C$, potem je*

$$g \circ f : A \to C.$$

Dokaz. Dokažimo najprej (i). Upoštevamo, da sta f in g enolični, dokazujemo pa enoličnost $g \circ f$.

$$\begin{aligned} a(g \circ f)b \text{ in } a(g \circ f)c &\sim \exists u\,(afu \text{ in } ugb) \text{ in } \exists v\,(afv \text{ in } vgc) \\ &\sim \exists u \exists v\,(afu \text{ in } ugb \text{ in } afv \text{ in } vgc) \\ &\Rightarrow \exists u \exists v\,(u = v \text{ in } ugb \text{ in } vgc) && (f \text{ je enolična}) \\ &\Rightarrow \exists u\,(ugb \text{ in } ugc) \\ &\Rightarrow b = c && (g \text{ je enolična}) \end{aligned}$$

Za drugi del trditve (i) računamo takole.

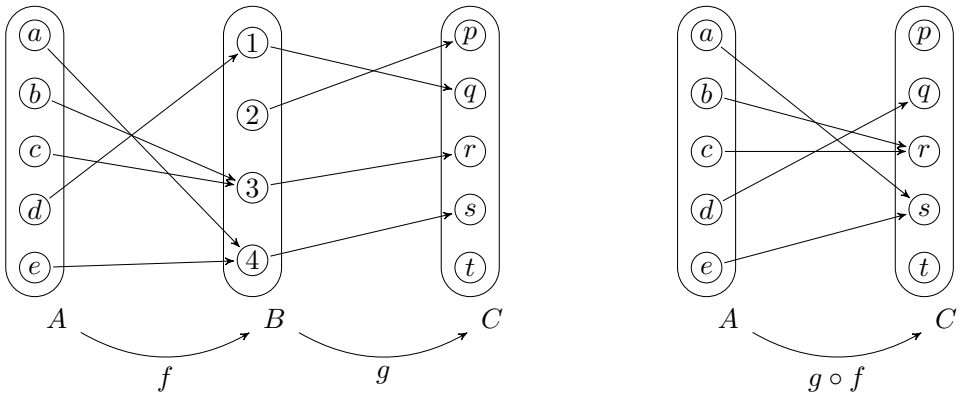
$$\begin{aligned} b = (g \circ f)(a) &\sim a(g \circ f)b \\ &\sim a(f * g)b \\ &\sim \exists u\,(afu \text{ in } ugb) \\ &\sim \exists u\,(u = f(a) \text{ in } b = g(u)) \\ &\sim b = g(f(a)) \end{aligned}$$

Za dokaz (ii) najprej upoštevamo, da je $g \circ f \subseteq A \times C$, saj je $g \subseteq B \times C$ in $f \subseteq A \times B$. Poleg tega je $\mathcal{Z}_f \subseteq B = \mathcal{D}_g$ in zato je $\mathcal{D}_{g \circ f} = \mathcal{D}_f = A$. Torej je definicijsko območje $g \circ f$ enako A, kar pomeni, da je $g \circ f$ preslikava iz A v C. □

Kompozitum preslikav $g \circ f$, kjer sta f in g iz (5.5), je prikazan na sliki 5.4.

Reciklirajmo na tem mestu nekatere od lastnosti relacijskega produkta, ki jih prepišemo v jezik kompozituma preslikav.

Trditev 5.3 *Naj bodo $f : A \to B$, $g : B \to C$ in $h : C \to D$. Potem velja*



Slika 5.4: Kompozitum preslikav $f : A \to B$ in $g : B \to C$ je preslikava $g \circ f : A \to C$.

(i) $f \circ \mathrm{id}_A = \mathrm{id}_B \circ f = f$.

(ii) *Kompozitum preslikav je asociativen,*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Dokaz. Kompozitum preslikave f z ustrezno identiteto lahko prepišemo v relacijski produkt, saj je

$$f \circ \mathrm{id}_A = \mathrm{id}_A * f.$$

Skličemo se lahko na trditev 4.1(v).

Asociativnost kompozituma preslikav pa prevedemo na asociativnost produkta relacij, trditev 4.1(iii),

$$h \circ (g \circ f) = (f * g) * h = f * (g * h) = (h \circ g) \circ f.$$

□

Mimogrede, kompozitum relacij ravno tako kot relacijski produkt ni komutativna operacija. V splošnem za par preslikav f, g kompozitum $f \circ g$ ni enak kompozitumu $g \circ f$.

Kaj dobimo, če komponiramo preslikavo f z (v splošnem relacijo) f^{-1}? Potrebno bo malo pazljivosti, saj f^{-1} morda ni enolična, kaj šele preslikava. Odgovor strnemo v naslednjo trditev.

Trditev 5.4 *Naj bo $f : A \to B$ in f^{-1} inverzna* relacija *k f. Z R_f označimo relacijo v množici A, za katero je xR_fy natanko tedaj, ko je $f(x) = f(y)$.*

(i) $f^{-1} \circ f = R_f$,

(ii) $f \circ f^{-1} = \mathrm{id}_{\mathcal{Z}_f}$,

(iii) *f je injektivna natanko tedaj, ko je* $f^{-1} \circ f = \mathrm{id}_A$,

(iv) *f je surjektivna natanko tedaj, ko je* $f \circ f^{-1} = \mathrm{id}_B$.

Dokaz. Dokazujemo po vrsti in pri tem upoštevamo, da f^{-1} morda ni enolična.

$$\begin{aligned} a(f^{-1} \circ f)b &\sim a(f * f^{-1})b \\ &\sim \exists u\,(afu \text{ in } uf^{-1}b) \\ &\sim \exists u\,(afu \text{ in } bfu) \\ &\sim \exists u\,(u = f(a) \text{ in } u = f(b)) \\ &\sim f(a) = f(b) \\ &\sim aR_f b \end{aligned}$$

Za (ii) postopamo podobno.

$$\begin{aligned} a(f \circ f^{-1})b &\sim a(f^{-1} * f)b \\ &\sim \exists u\,(af^{-1}u \text{ in } ufb) \\ &\sim \exists u\,(ufa \text{ in } ufb) \\ &\sim \exists u\,(a = f(u) \text{ in } b = f(u)) \\ &\sim a = b \text{ in } a \in \mathcal{Z}_f \text{ in } b \in \mathcal{Z}_f \\ &\sim a\,\mathrm{id}_{\mathcal{Z}_f} b \end{aligned}$$

V dokazih (iii) in (iv) seveda uporabimo lastnosti (i) in (ii).

$$\begin{aligned} f \text{ je injektivna} &\sim \forall x \forall y\,(f(x) = f(y) \Leftrightarrow x = y) \\ &\sim \forall x \forall y\,(xR_f y \Leftrightarrow x\,\mathrm{id}_A y) \\ &\sim R_f = \mathrm{id}_A \\ &\sim f^{-1} \circ f = \mathrm{id}_A \end{aligned}$$

$$\begin{aligned} f \text{ je surjektivna} &\sim \mathcal{Z}_f = B \\ &\sim \mathrm{id}_{\mathcal{Z}_f} = \mathrm{id}_B \\ &\sim f \circ f^{-1} = \mathrm{id}_B \end{aligned}$$

□

Posledica 5.5 *Naj bo $f : A \to B$ bijekcija. Potem je inverzna preslikava $f^{-1} : B \to A$ bijekcija in velja*

$$f \circ f^{-1} = \mathrm{id}_B \quad \textit{in} \quad f^{-1} \circ f = \mathrm{id}_A. \tag{5.7}$$

Dokaz. Enakosti (5.7) sta enostavni posledici točk (iii) in (iv) trditve 5.4.

Potrebno je samo premisliti, da je inverzna preslikava f^{-1} res bijekcija. Upoštevamo, da je $(f^{-1})^{-1} = f$. Zato je

$$(f^{-1})^{-1} \circ f^{-1} = f \circ f^{-1} = \mathrm{id}_B.$$

Če trditev 5.4(iii) uporabimo na preslikavi $f^{-1} : B \to A$, pridelamo, da je f^{-1} injektivna.

Če na preslikavi $f^{-1} : B \to A$ uporabimo tudi trditev 5.4(iv), pridelamo tudi surjektivnost iste preslikave f^{-1}, saj je po predpostavki f injektivna, torej velja $f^{-1} \circ f = \mathrm{id}_A$. □

Razdelek zaključimo s tehnično trditvijo. Obstoj surjektivne preslikave med množicama A in B bomo izenačili z obstojem injektivne preslikave med istima množicama *v drugi smeri*.

Trditev 5.6 *Naj bosta A in B neprazni množici.*

(i) *Če obstaja injektivna preslikava $f : A \to B$, potem obstaja tudi surjektivna preslikava $\bar{f} : B \to A$.*

(ii) *Če obstaja surjektivna preslikava $g : A \to B$, potem obstaja tudi injektivna preslikava $\bar{g} : B \to A$.*[1]

Dokaz. Obakrat, tako v (i) kot v (ii), bomo privzeli, da f oziroma g nista bijekciji. Sicer lahko shajamo kar z inverzno preslikavo.

Naj bo torej $f : A \to B$ injektivna preslikava, ki ni surjektivna. Po izreku 5.1 je f^{-1} enolična, obenem pa je njeno definicijsko območje $\mathcal{D}_{f^{-1}}$ prava podmnožica B. Izberimo poljuben element $a \in A$ in razširimo f^{-1} do preslikave $\bar{f} : B \to A$ z naslednjim opisom:

$$\bar{f}(b) = \begin{cases} f^{-1}(b), & b \in \mathcal{D}_{f^{-1}} \\ a, & b \in B \setminus \mathcal{D}_{f^{-1}} \end{cases} \tag{5.8}$$

Ker je $\mathcal{Z}_{f^{-1}} = \mathcal{D}_f = A$ in je $\bar{f}$ razširitev f^{-1}, je $\bar{f} : B \to A$ surjektivna preslikava. Mimogrede, tako definirana preslikava $\bar{f}$ ni injektivna, saj ima vsak element množice $B \setminus \mathcal{D}_{f^{-1}}$ (vsaj eden obstaja) s preslikavo $\bar{f}$ isto funkcijsko vrednost kot $f(a)$.

Za (ii) je potrebno nekoliko več tehničnih podrobnosti. Za vsak $b \in B$ lahko definiramo *vrečo*

$$A_b = \{a \in A \mid g(a) = b\}.$$

Ker je g surjektivna preslikava, so množice A_b neprazne in lahko iz vsake od njih *izberemo* predstavnika $a_b \in A_b$. Injektivno preslikavo $\bar{g} : B \to A$ definiramo z naslednjim opisom:

$$\bar{g}(b) = a_b \in A_b$$

[1] Potrebujemo *aksiom izbire*. Le-tega v rigorozni obravnavi teorije množic velikokrat enostavno privzamemo. V naivni obravnavi teorije množic pa ga, skupaj s preostalimi, manj kontroverznimi, aksiomi modro zamolčimo.

Preslikava $\bar{g}$ je dobro definirana. Če je $b \neq b'$, potem sta množici A_b in $A_{b'}$ disjunktni. Za element a^* iz preseka $A_b \cap A_{b'}$ bi namreč veljalo $f(a^*) = b$ (saj $a^* \in A_b$) in $f(a^*) = b'$ (saj $a^* \in A_{b'}$), kar je nemogoče. Posledično sta predstavnika $a_b \in A_b$ in $a_{b'} \in A_{b'}$ različna in zato je preslikava $\bar{g}$ injektivna. Tudi v tem primeru konstruirana preslikava $\bar{g}$ ni surjektivna, saj ni vsak element $a \in A$ predstavnik katere od množic A_b (ker preslikava g ni injektivna, ima namreč vsaj ena od vreč vsaj dva elementa). □

5.3 Lastnosti preslikav in kompozitum

V tem razdelku obravnavamo, kako se injektivnost in surjektivnost preslikav obnašata, če preslikave komponiramo. Ni se težko prepričati, da je družina injektivnih preslikav zaprta za komponiranje, ravno tako družina surjektivnih preslikav.

Izrek 5.7 *Izberimo poljubni preslikavi $f : A \to B$ in $g : B \to C$.*

(i) *Če sta f in g injektivni, potem je tudi preslikava $g \circ f : A \to C$ injektivna.*

(ii) *Če sta f in g surjektivni, potem je tudi preslikava $g \circ f : A \to C$ surjektivna.*

(iii) *Če sta f in g bijektivni, potem je tudi preslikava $g \circ f : A \to C$ bijektivna.*

Dokaz. Privzemimo, da sta preslikavi f in g injektivni, in dokazujmo (i).

Izberimo poljubna različna elementa $a \neq a'$, za katera velja $a, a' \in A$. Označimo $b = f(a)$ in $b' = f(a')$. Ker je f injektivna, sta b in b' različna elementa množice B.

Vstavimo ju v preslikavo g in označimo $c = g(b)$ in $c' = g(b')$. Recikliramo sklep. Tudi preslikava g je injektivna, zato imata različna elemeta b, b' različni sliki c, c'.

Na koncu upoštevajmo, da je $c = g(f(a)) = (g \circ f)(a)$ in $c' = g(f(a')) = (g \circ f)(a')$. Preslikava $g \circ f$ vsaka dva različna elementa $a \neq a'$ preslika v različni sliki. Zato je injektivna.

Izberimo zdaj poljuben $c \in C$. Za dokaz surjektivnosti preslikave $g \circ f$ je dovolj poiskati element množice A, ki ga preslikava $g \circ f$ slika v c. Pri tem bomo seveda privzeli, da sta tako f kot g surjekciji.

Ker je preslikava g surjektivna, element c pripada njeni sliki. Zato obstaja $b \in B$, za katerega je $c = g(b)$. Tudi preslikava f je surjektivna, zato obstaja $a \in A$, za katerega je $b = f(a)$.

Torej velja $c = g(f(a)) = (g \circ f)(a)$ in dokaz (ii) je zaključen.

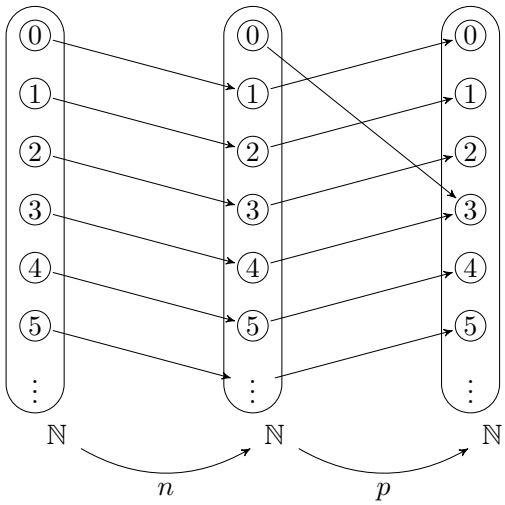
Točka (iii) je direktna posledica (i) in (ii). □

Velja morda obrat izreka 5.7 — ali lahko iz injektivnosti oziroma surjektivnosti kompozituma $g \circ f$ sklepamo, da sta preslikavi f in g injektivni oziroma surjektivni?

V splošnem je odgovor negativen. Definirajmo preslikavo $n : \mathbb{N} \to \mathbb{N}$ s predpisom $n(x) = x + 1$, preslikava n naravnemu številu določi naslednika. Preslikava n ni surjektivna, saj 0 ni naslednik nobenega naravnega števila.

Podobno definirajmo tudi preslikavo $p : \mathbb{N} \to \mathbb{N}$, ki naravno število 0 preslika v, denimo, 3, večja naravna števila pa slika v njihove predhodnike. Ne glede na to, kako bi definirali funkcijsko vrednost ničle 0, preslikava p ne bo injektivna.

Na sliki 5.5 je prikazan kompozitum preslikav $p \circ n$. Enostavno se je prepričati, da je kompozitum $p \circ n = \mathrm{id}_{\mathbb{N}}$, ki je celo bijektivna preslikava.



Slika 5.5: Kompozitum preslikav $p \circ n$ je lahko bijekcija, četudi n ni surjektivna in p ni injektivna preslikava.

Delni sklep vseeno velja.

Izrek 5.8 *Naj bo $f : A \to B$ in $g : B \to C$.*

(i) *Če je $g \circ f : A \to C$ injektivna, potem je injektivna tudi preslikava f.*

(ii) *Če je $g \circ f : A \to C$ surjektivna, potem je surjektivna tudi preslikava g.*

Dokaz. Dokazovali bomo kontrapoziciji zgornjih implikacij.

Pri (i) privzamemo, da f ni injektivna. To pomeni, da obstajata različna elementa $a \neq a'$, za katera je $f(a) = f(a')$. Odtod sledi, da je $g(f(a)) = g(f(a'))$ oziroma

$$(g \circ f)(a) = (g \circ f)(a')$$

To pomeni, da $g \circ f$ ni injektivna.

Če preslikava g ni surjektivna, potem obstaja $c \in C \setminus \mathcal{Z}_g$. Ker je $\mathcal{Z}_{g\circ f} \subseteq \mathcal{Z}_g$, element c ne pripada zalogi vrednosti preslikave $g \circ f$. Zato $g \circ f$ ni surjektivna. □

Razdelek zaključimo s kategorično karakterizacijo inverzne preslikave. Če je $f : A \to B$ bijekcija, je po posledici 5.5 tudi $f^{-1} : B \to A$ bijektivna preslikava. Kompozitum preslikave f in njej inverzne preslikave f^{-1} pa je ustrezna identiteta, odvisno od vrstnega reda preslikav v operaciji.

Naslednji izrek dokazuje tudi obrat. Inverzno preslikavo bi lahko definirali z opisom, kaj takšna preslikava počne.

Izrek 5.9 *Naj bosta $f : A \to B$ in $g : B \to A$. Če veljata zvezi*

$$f \circ g = \mathrm{id}_B \quad in \quad g \circ f = \mathrm{id}_A,$$

potem sta preslikavi f in g bijektivni in velja $f^{-1} = g$.

Dokaz. Upoštevamo, da sta preslikavi id_A in id_B tako surjektivni kot injektivni.

Iz $g \circ f = \mathrm{id}_A$ z uporabo izreka 5.8 pridelamo, da je f injektivna in g surjektivna. Iz $f \circ g = \mathrm{id}_B$ z uporabo istega izreka pridelamo, da je g injektivna in f surjektivna. Torej sta tako f kot g bijekciji.

Po posledici 5.5 sta tako f^{-1} kot g^{-1} bijektivni preslikavi.

Računajmo

$$f^{-1} = f^{-1} \circ \mathrm{id}_B = f^{-1} \circ (f \circ g) = (f^{-1} \circ f) \circ g = \mathrm{id}_A \circ g = g$$

Dokaz je s tem zaključen. □

5.4 Slike in praslike

Preslikavo $f : A \to B$ lahko na naraven način *raztegnemo* na podmnožice množice A. Pri $A' \subseteq A$ bi radi vedeli, kaj so slike elementov iz A'. Zanima nas torej množica

$$\{f(a) \mid a \in A'\}.$$

Pri izbrani $B' \subseteq B$ pa želimo vedeti, kateri elementi množice A imajo sliko v B', kaj je

$$\{a \mid f(a) \in B'\}$$

Torej smemo definirati preslikavi

$$\begin{aligned} &F \,:\, \mathcal{P}A \to \mathcal{P}B, \\ &F(A') \,=\, \{f(a')\,;\, a' \in A'\} \subseteq B\,, \end{aligned} \tag{5.9}$$

in

$$\begin{aligned} &F^{-1} \,:\, \mathcal{P}B \to \mathcal{P}A, \\ &F^{-1}(B') \,=\, \{a'\,;\, a' \in A \,\wedge f(a) \in B'\} \subseteq A\,. \end{aligned} \tag{5.10}$$

Preslikavo F imenujemo tudi *slika*, preslikavo F^{-1} pa *praslika* glede na originalno preslikavo $f : A \to B$.

Na mestu je komentar. V literaturi za preslikavo slike in praslike največkrat uporabljamo kar oznaki f in f^{-1}, potem pa iz konteksta razberemo, ali simbol f^{-1} dejansko pomeni inverzno preslikavo k f (če le-ta sploh obstaja) ali pa morda prasliko.

Četudi se bomo v zapisu z uporabo velikih črk F, F^{-1} izognili dvoumnemu pomenu, preslikava praslike F^{-1} v splošnem *ni* inverzna preslikava k preslikavi slike F. V nadaljevanju bomo poiskali tudi natančen pogoj, kdaj sta slika F in praslika F^{-1} paroma inverzni preslikavi.

Kakšne so lastnosti preslikav F in F^{-1}? Če kot argument slike F vstavimo singleton $\{a\}$, potem ne dobimo veliko novega, $F(\{a\}) = \{f(a)\}$.

Naj bosta $A', A'' \subseteq A$.

$$\begin{aligned} F(A') \cup F(A'') &= \{f(a) \mid a \in A'\} \cup \{f(a) \mid a \in A''\} \\ &= \{f(a) \mid a \in A' \ \text{ ali } \ a \in A''\} \\ &= F(A' \cup A'') \end{aligned}$$

Analogna formula za presek ne velja. Protiprimer pridelamo s preslikavo f, ki ni injektivna. Če, na primer, za različna $a' \neq a''$ velja $f(a') = f(a'')$, potem je po eni strani $F(\{a'\}) \cap F(\{a''\}) = \{f(a')\} \neq \emptyset$, po drugi strani pa je $F(\{a'\} \cap \{a''\}) = F(\emptyset) = \emptyset$.

Praslika se v tem pogledu obnaša lepše. Naj bosta $B', B'' \subseteq B$.

$$\begin{aligned} F^{-1}(B') \cup F^{-1}(B'') &= \{a \mid f(a) \in B'\} \cup \{a \mid f(a) \in B''\} \\ &= \{a \mid f(a) \in B' \ \text{ ali } \ f(a) \in B''\} \\ &= \{a \mid f(a) \in B' \cup B''\} \\ &= F^{-1}(B' \cup B'') \end{aligned}$$

Račun za presek poteka analogno.

$$\begin{aligned} F^{-1}(B') \cap F^{-1}(B'') &= \{a \mid f(a) \in B'\} \cap \{a \mid f(a) \in B''\} \\ &= \{a \mid f(a) \in B' \ \text{ in } \ f(a) \in B''\} \\ &= \{a \mid f(a) \in B' \cap B''\} \\ &= F^{-1}(B' \cap B'') \end{aligned}$$

Druga enakost, ki ne velja v primeru slike, pri prasliki ne povzroča težav.

Zgornje račune strnemo v naslednjo trditev.

Trditev 5.10 *Naj bo $f : A \to B$ preslikava, F oziroma F^{-1} ustrezni preslikavi slike oziroma praslike, $A', A'' \subseteq A$ ter $B', B'' \subseteq B$. Potem je*

(i) $F(A' \cup A'') = F(A') \cup F(A'')$

(ii) $F(A' \cap A'') \subseteq F(A') \cap F(A'')$

(iii) $F^{-1}(B' \cup B'') = F^{-1}(B') \cup F^{-1}(B'')$

(iv) $F^{-1}(B' \cap B'') = F^{-1}(B') \cap F^{-1}(B'')$

Dokaz. Potrebujemo samo še utemeljitev (ii). Izberimo poljuben $b \in F(A' \cap A'')$. Po definiciji slike F obstaja $a \in A' \cap A''$, za katerega je $b = f(a)$. V tem primeru $b \in F(A')$ in tudi $b \in F(A'')$, zato $b \in F(A') \cap F(A'')$. □

Navedimo še karakterizacijo lastnosti preslikav slike F in praslike F^{-1}.

Trditev 5.11 *Naj bo $f : A \to B$ preslikava in F oziroma F^{-1} ustrezni preslikavi slike oziroma praslike. Potem je*

(i) slika F *injektivna natanko tedaj, ko je f injektivna,*

(ii) slika F *surjektivna natanko tedaj, ko je f surjektivna,*

(iii) praslika F^{-1} *injektivna natanko tedaj, ko je f surjektivna, in*

(iv) praslika F^{-1} *surjektivna natanko tedaj, ko je f injektivna.*

Dokaz. Za (i) najprej privzemimo injektivnost slike F in izberimo poljubna različna elementa $a, a' \in A$. Ker je F po predpostavki injektivna, je $\{f(a)\} = F(\{a\}) \neq F(\{a'\}) = \{f(a')\}$. Zato je tudi $f(a) \neq f(a')$.

Privzemimo zdaj injektivnost preslikave f in izberimo različni podmnožici $A' \neq A''$ množice A. Ker je $A' \neq A''$, obstaja $a \in A' + A''$. Za $f(a)$ velja, da pripada natančno eni od množic $F(A'), F(A'')$ (Če, denimo, $a \in A'$, potem iz $f(a) \in A''$ sledi, da obstaja $a'' \in A''$, za katerega je $f(a'') = f(a)$. To je nemogoče ob injektivni preslikavi f). Zato sta $F(A')$ in $F(A'')$ različni.

Če je F surjektivna preslikava, potem je tudi vsak singleton $\{b\} \subseteq B$ v sliki preslikave F. Z drugimi besedami, pri poljubnem $b \in B$ obstaja množica $A' \subseteq A$, za katero je $F(A') = \{b\}$. Za vsak $a \in A'$ torej velja $f(a) = b$, kar posledično pomeni, da $b \in \mathcal{Z}_f$. Ker je b v osnovi poljuben, je f surjektivna.

Za drugo smer implikacije izberimo poljubno podmnožico $B' \subseteq B$. Za vsak $b \in B'$ lahko zaradi surjektivnosti preslikave f poiščemo original a, za katerega je $b = f(a)$. Družina vseh takšnih originalov se z F preslika natanko v B'.

Dokaza točk (iii) in (iv) izdelamo s podobnimi premisleki, zato ju na tem mestu izpustimo. □

V začetku razdelka smo opozorili, da preslikavi slike in praslike, F in F^{-1}, nista nujno paroma inverzni. Za konec razdelka in poglavja odgovorimo tudi na vprašanje, kdaj pa se to vseeno zgodi.

Trditev 5.12 *Naj bo $f : A \to B$ preslikava in F oziroma F^{-1} ustrezni preslikavi slike oziroma praslike. Potem je F^{-1} inverzna preslikava k F natanko tedaj, ko je f bijekcija.*

Dokaz. Dokaz ni težaven. Če sta F in F^{-1} paroma inverzni preslikavi (to mimogrede, glej izrek 5.1(ii) in posledico 5.5, pomeni, da sta bijekciji), potem je po trditvi 5.11 preslikava f bijekcija.

Po drugi strani, če je f bijekcija, potem sta po trditvi 5.11 preslikavi slike F in praslike F^{-1} bijektivni. V tem primeru za vsako množico $A' \subseteq A$ velja $F^{-1}(F(A)) = A$, za vsako množico $B' \subseteq B$ pa $F(F^{-1}(B)) = B$. Torej sta, skličemo se na izrek 5.9, F in F^{-1} res paroma inverzni preslikavi. □

Poglavje 6

Moč končnih množic

Štetje objektov v množici se zdi prva nadgradnja operacij z množicami. Če smo absolvirali unijo, presek, simetrično razliko (in še nekaj bolj zapletenih operacij), se lahko posvetimo tudi preštevanju elementov v množici.

6.1 Končne množice

Poskusimo na začetku odgovoriti na vprašanje, kaj sploh je končna množica, oziroma, kako ločiti med končnimi in neskončnimi množicami.

Zdi se, da je naloga kar se da enostavna. Takoj ko množico vidimo, vemo, katerega tipa je. Tako je množica črk slovenske abecede seveda končna, ravno tako množica ničel kvadratne enačbe $x^2 + x - 1 = 0$. Po drugi strani pa je množica naravnih števil $\mathbb{N}$ neskončna, da o množici realnih števil $\mathbb{R}$ sploh ne govorimo.

Množica črk slovenske abecede ima tako natanko 25 črk, množica rešitev izbrane kvadratne enačbe pa ima med nič in dvema elementoma. Vsekakor ne preveč.

Zdi se, da za definicijo končnih (in posredno tudi neskončnih množic) potrebujemo pojem naravnega števila. Množica A je *končna* natanko tedaj, ko je število njenih elementov enako kateremu od naravnih števil $n \in \mathbb{N}$. Množice, ki niso končne, pa so *neskončne*.

To je ena možnost.

Alternativni pristop zmore definirati pojem končne množice brez uporabe naravnih števil, je pa potrebno zgodbo začeti na koncu.

Začnimo takole. Za množici A in B pravimo, da sta *enako močni*, če med njima obstaja bijektivna preslikava

$$f^{\text{bijektivna}} : A \to B.$$

Pišemo tudi $|A| = |B|$.

Preslikava identitete je bijektivna, zato za vsako množico A velja $|A| = |A|$. Ravno tako

iz $|A| = |B|$ sledi $|B| = |A|$, saj je za bijektivno preslikavo $f : A \to B$ tudi njej inverzna preslikava $f^{-1} : B \to A$ bijekcija.

Končno iz $|A| = |B|$ in $|B| = |C|$ sledi $|A| = |C|$, saj je kompozitum bijekcij $A \to B$ in $B \to C$ bijekcija med A in C, glej izrek 5.7.

Množica A je *neskončna*, če obstaja kakšna njena prava podmnožica $A' \subset A$, za katero je $|A'| = |A|$. Množica B je *končna*, če ni neskončna. To pomeni, da pri nobeni pravi podmnožici $B' \subset B$ ne obstaja bijektivna preslikava med B' in B.

Na srečo se obe definiciji končnih in neskončnih množic, prva intuitivna in zadnja z bijektivnimi preslikavami, ujemata.[1]

Množica naravnih števil $\mathbb{N}$ je neskončna. Po eni strani zato, ker za vsako naravno število n velja, da ima $\mathbb{N}$ strogo več elementov kot n. Z alternativno definicijo pa zato, ker obstaja bijektivna preslikava med $\mathbb{N}$ in njeno pravo podmnožico sodih števil $\mathbb{S} = \{0, 2, 4, \ldots\}$, takšno preslikavo lahko predstavimo z opisom

$$n \mapsto 2n.$$

Množica $B = \{\heartsuit,\diamondsuit,\clubsuit,\spadesuit\}$ je končna. Vsaka prava podmnožica B' množice B ima največ tri elemente. Če je $\varphi : B' \to B$ injektivna preslikava, potem ima zaloga vrednosti $\mathcal{Z}_\varphi$ natanko toliko elementov kot B', vsekakor strogo manj od 4. Zato φ ni bijekcija.

Naj med množico A in njeno pravo podmnožico $A' \subset A$ obstaja bijektivna preslikava

$$\varphi : A' \to A.$$

Množica A je seveda neskončna. Inverzni preslikavi $\varphi^{-1} : A \to A'$ razširimo kodomeno do cele množice A in

$$\varphi^{-1} : A \to A$$

je preslikava množice A *v isto množico* A, ki je *injektivna* in *ni surjektivna*. Z izbiro poljubnega elementa $\bar{a} \in A$ pa lahko preslikavo $\varphi : A' \to A$ razširimo do *surjektivne* in *ne injektivne* preslikave $\bar{\varphi} : A \to A$ z opisom

$$\bar{\varphi}(a) = \begin{cases} \varphi(a), & a \in A', \\ \bar{a}, & a \in A \setminus A'. \end{cases}$$

Pokazali smo naslednjo lastnost. Če je A neskončna množica, potem za preslikave $A \to A$ lastnosti injektivnosti in surjektivnosti nista enakovredni. Za vsako *neskončno* množico A obstaja preslikava $A \to A$, ki je samo surjektivna in ni injektivna. Ravno tako pa obstaja preslikava $A \to A$, ki je injektivna in ni surjektivna.

Za končne množice je zgodba drugačna. Velja namreč Dirichletov princip[2].

[1] Tudi za to trditev potrebujemo *aksiom izbire*.

[2] V preštevalni matematiki pojem Dirichletov princip ponavadi uporabljamo za dejstvo, da ne obstaja injektivna preslikava iz ene končne množice A v drugo množico B, če ima množica B strogo manjšo moč kot A.

Izrek 6.1 (Dirichletov princip) *Naj bo B končna množica in $\psi : B \to B$ preslikava končne množice B vase. Naslednje trditve so enakovredne:*

(i) *ψ je surjektivna.*

(ii) *ψ je injektivna.*

(iii) *ψ je bijektivna.*

Dokaz. Če je $\psi : B \to B$ bijekcija, je seveda tudi surjekcija in injekcija hkrati. Bistvo izreka je enakovrednost med (i) in (ii).

Če za preslikave iz $B \to B$ lastnosti injektivnosti in surjektivnosti nista enakovredni, potem obstaja preslikava $B \to B$, ki ima natančno eno od obeh lastosti. Po izreku 5.6 lahko privzamemo, da obstaja preslikava $\psi : B \to B$, ki je injektivna in ni surjektivna.

Potem je ψ^{-1} *bijektivna* preslikava iz $B' = \mathcal{Z}_f$ v B. To pomeni, da obstaja bijektivna preslikava med pravo podmnožico $B' \subset B$ in celo množico B. To je v protislovju s predpostavko, da je B končna množica. □

Na tem mestu lahko *moč končne množice* enačimo s številom njenih elementov. Tako velja

$$|\{a, b, c\}| = 3 \quad \text{in} \quad |\{1, 2\}| = 2.$$

Prazna množica je edina množica z natanko 0 elementi, moč množice pa je neodvisna od strukture njenih elementov. Torej je

$$|\emptyset| = 0 \quad \text{in} \quad |\{\emptyset\}| = |\{\mathbb{N}\}| = 1.$$

Definirajmo kanonično končno množico z natanko $n \in \mathbb{N}$ elementi

$$[n] = \{0, \ldots, n-1\}, \quad n \in \mathbb{N}, \tag{6.1}$$

in kanonično družino preslikav

$$\varphi_{m,n} : [m] \to [n], \quad m, n \in \mathbb{N} \setminus \{0\}, \tag{6.2}$$

definirano z opisom

$$\varphi_{m,n}(x) = \min(\{x, n-1\}). \tag{6.3}$$

Preslikava $\varphi_{m,n}$ je injektivna natanko tedaj, ko je $m \leq n$, in surjektivna natanko tedaj, ko je $m \geq n$ (in bijektivna natanko tedaj, ko je $m = n$).

6.2 Moč končnih množic in operacije

V tem razdelku bomo raziskali zveze med operacijami z množicami in številom elementov. Presenetljivo bo, da bodo formule za moč množic bolj enostavne v primeru bolj zapletenih operacij z množicami.

V preostanku razdelka naj bosta A in B končni množici, njuni moči pa označimo z $a = |A|$ in $b = |B|$.

(1) Kartezični produkt:

$$|A \times B| = |A| \cdot |B| \tag{6.4}$$

(2) Družina preslikav:

$$|A^B| = |A|^{|B|} \tag{6.5}$$

(3) Potenčna množica:

$$|\mathcal{P}A| = 2^{|A|} \tag{6.6}$$

Za utemeljitve (1), (2) in (3) se lahko omejimo na primer, ko je $A = [a]$ in $B = [b]$. Sicer najprej poiščemo (v dokazu (1), pri (2) in (3) uporabimo analogen trik) bijekcijo med $A \times B$ in $[a] \times [b]$.

Za (1) se najprej lotimo primera, ko je katera izmed množic A oziroma B prazna. V tem primeru je prazen tudi kartezični produkt $A \times B$ in sta tako $|A \times B|$ kot $|A| \cdot |B|$ enaka 0.

V nadaljevanju se lahko omejimo na neprazni množici A in B. Zdaj je dovolj pokazati, da obstaja bijektivna preslikava med $[a] \times [b]$ in $[ab]$. Ni je težko konstruirati, primer takšne bijekcije je preslikava

$$\zeta : (p, q) \mapsto bp + q.$$

Ker je $p < a$ in $q < b$, je $\zeta(a, b) = bp + q \leq b(a-1) + (b-1) = ab - 1 < ab$. Preslikava ζ je torej dobro definirana.

Denimo, da je $\zeta(p, q) = \zeta(p', q')$. Torej velja enakost

$$bp + q = bp' + q'.$$

S preoblikovanjem lahko zgornjo enakost prepišemo v

$$b(p - p') = q' - q.$$

Če sta obe strani zgornje enačbe enaki 0, potem velja $(p, q) = (p', q')$, saj je $b \neq 0$. V nasprotnem primeru sta obe strani enačbe od 0 različni, poleg tega pa je izraz $q' - q$ deljiv s številom b. Ker sta tako q' kot q strogo manjša od b, je $q' - q$ po absolutni vrednosti strogo manjši od b in različen od 0. To ni mogoče in zato je preslikava ζ injektivna.

Po drugi strani pa je preslikava ζ surjektivna, saj lahko vsako naravno število iz $[ab]$ (celo na enoličen način) zapišemo v obliki $bp + q$ pri ustreznih $p \in [a]$ in $q \in [b]$.

Pri (2) označimo z f preslikavo iz $[b]$ v $[a]$. Preslikavo f natanko opišemo s funkcijskimi vrednostmi za vsak $q \in [b]$. Za izbor funkcijske vrednosti $f(q)$ imamo natanko a možnosti, za opis celotne preslikave f je potrebno izbiro ponoviti b-krat. Torej je vseh preslikav v A^B natanko a^b.

Vsako podmnožico A' množice $[a]$ lahko opišemo s karakteristično funkcijo $\chi_{A'} : [a] \to \{0, 1\}$, definirano z opisom

$$\chi_{A'}(p) = \begin{cases} 1, & p \in A', \\ 0, & p \in [a] \setminus A'. \end{cases}$$

In obratno. Karakteristične funkcije so v bijektivni zvezi s podmnožicami. Število karakterističnih funkcij $[a] \to \{0, 1\}$ pa je po predhodni točki (2) enako 2^a. Posledično je $|\mathcal{P}A| = 2^a$.

Unija, presek, razlika množic

Unijo, presek in razliko množic uženemo z naslednjim principom. Denimo, da želimo preveriti veljavnost formule

$$|A| = |B| + |C| - |D|.$$

V tem primeru je potrebno pokazati, da je prispevek vsakega elementa $a \in A$ tudi v izrazu na desni strani natanko 1. Če pa $a \notin A$, potem mora biti njegov skupen prispevek na desni enak 0.

Denimo, da a pripada izključno množici B. Prispevek takšnega elementa v množici B je enak 1, v množicah C in D pa nič. Skupen prispevek elementa a na desni strani bi v tem primeru bil enak 1.

Kot v predhodnem razdelku bomo tudi tukaj privzeli, da so množice A, B in C končne.

(4) Razlika množic:
če je $B \subseteq A$, potem je $|A \setminus B| = |A| - |B|$. V splošnem je

$$|A \setminus B| = |A| - |A \cap B| \tag{6.7}$$

(5) Unija dveh množic:
če je $A \cap B = \emptyset$, potem je $|A \cup B| = |A| + |B|$. V splošnem je

$$|A \cup B| = |A| + |B| - |A \cap B| \tag{6.8}$$

(6) Unija treh množic:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C| \tag{6.9}$$

Za posebno varianto (4) je dovolj opaziti dvoje. Če $a \in A \setminus B$, potem je njegov prispevek na desni v izrazu $|A| - |B|$ enak 1, saj $a \in A$ in $a \notin B$. Če pa $a \notin A \setminus B$, potem je njegov prispevek na desni enak 0, saj a bodisi pripada obema ali pa nobeni izmed množic A, B.

Za splošno verzijo (4) računajmo takole.

$$\begin{aligned} A \setminus (A \cap B) &= A \cap (A \cap B)^c = A \cap (A^c \cup B^c) \\ &= (A \cap A^c) \cup (A \cap B^c) = \emptyset \cup (A \cap B^c) = A \cap B^c = A \setminus B \end{aligned}$$

Zato je $|A \setminus B| = |A \setminus (A \cap B)|$, pri čemer je $A \cap B \subseteq A$.

Za (5) ločimo dva primera. Če $a \in A \cup B$, potem v vsaj eni od množic A oziroma B prispeva enico. V primeru, ko je $A \cap B = \emptyset$, prispevki nobenega elementa $a \in A$ niso enaki 2.

Če pa je $A \cap B \neq \emptyset$, moramo elemente a iz preseka $A \cap B$ obravnavati posebej. Vsak takšen element v formuli (6.8) prispeva natanko $1 + 1 - 1 = 1$.

Za (6) ločimo primere glede na število množic (izmed A, B, C), katerim pripada vsak posamezen element a iz unije $A \cup B \cup C$. Če je to število enako 1, potem a prispeva enico v natanko enem od členov $|A|, |B|, |C|$, pri vseh ostalih členih pa je njegov prispevek enak 0.

Če a pripada natanko dvema od množic izmed A, B, C, potem je njegov skupni prispevek $1 + 1 - 1$, saj pripada natanko enemu od dvojnih presekov. Če pa a pripada $A \cap B \cap C$, je njegov prispevek enak $1 + 1 + 1 - 1 - 1 - 1 + 1$, kar skupaj znese natanko 1.

Princip vključitve in izključitve

Formuli (5) in (6) za izračun moči unije imata skupno rdečo nit. Od vsote moči posameznih množic odštejemo moči vseh dvojnih presekov (morda je en sam) in prištejemo moči vseh trojnih presekov (tak je kvečjemu en). Formulo lahko posplošimo tudi na unije z večjim številom členov.

Trditev 6.2 (princip vključitve in izključitve) *Naj bo $A_1, A_2, \ldots, A_n$ zaporedje n končnih množic. Potem je*

$$\begin{aligned} |A_1 \cup A_2 \cup \ldots \cup A_n| = {} & + |A_1| + |A_2| + \cdots + |A_n| \\ & - |A_1 \cap A_2| - |A_1 \cap A_3| - \cdots - |A_{n-1} \cap A_n| \\ & + |A_1 \cap A_2 \cap A_3| + \cdots + |A_{n-2} \cap A_{n-1} \cap A_n| \\ & \cdots \\ & + (-1)^{n+1} |A_1 \cap A_2 \cap \ldots \cup A_n|. \end{aligned}$$

V strnjeni obliki lahko zgornjo formulo prepišemo kot

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = \sum_{k=1}^{n} \left((-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} |A_{i_1} \cap \ldots \cap A_{i_k}| \right). \tag{6.10}$$

Dokaz. Celoten preostanek razdelka bomo namenili dokazu trditve 6.2.

Na začetku se prepričamo, da je v primeru $n = 2$ oziroma $n = 3$ formula (6.10) enakovredna formulama (6.8) in (6.9) za izračun moči unije dveh oziroma treh množic, ki smo ju spoznali v prejšnjem razdelku.

Binomski izrek trdi

$$(x+1)^m = \sum_{k=0}^{m} \binom{m}{k} x^k. \tag{6.11}$$

V (6.11) vstavimo $x = -1$. Vrednost izraza bo tipično 0, v primeru eksponenta $m = 0$ pa vseeno[3] 1. S preoblikovanjem pridelamo naslednjo formulo:

$$\sum_{k=1}^{m} (-1)^k \binom{m}{k} = \begin{cases} -1, & m > 0, \\ 0, & m = 0. \end{cases} \tag{6.12}$$

Izberimo poljuben element a. Število množic v zaporedju $A_1, A_2, \ldots, A_n$, ki vsebujejo element a, označimo z ℓ. Očitno velja $0 \le \ell \le n$. Izračunali bomo prispevek elementa a v desni strani enakosti (6.10).

Fiksirajmo k in izračunajmo, kakšen je prispevek elementa a v izrazu

$$\sum_{1 \le i_1 < \cdots < i_k \le n} |A_{i_1} \cap \ldots \cap A_{i_k}|. \tag{6.13}$$

Če je $\ell < k$ potem je prispevek a v (6.13) enak 0, saj a v tem primeru ne pripada nobenemu od presekov po k množic. Če pa je $\ell \ge k$, potem a v izrazu 6.13 prispeva $\binom{\ell}{k}$, saj preseke po k množic, ki vsebujejo element a, dobimo tako, da jih izberemo med tistimi ℓ množicami, ki vsebujejo a.

Zato je prispevek elementa a v izrazu

$$\sum_{k=1}^{n} \left((-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} |A_{i_1} \cap \ldots \cap A_{i_k}| \right)$$

enak

$$\sum_{k=1}^{n} (-1)^{k+1} \binom{\ell}{k} = -\sum_{k=1}^{\ell} (-1)^k \binom{\ell}{k} = \begin{cases} 1, & \ell > 0, \\ 0, & \ell = 0, \end{cases}$$

[3] V diskretni matematiki je 0^0 enako 1. S takšno izbiro dosežemo, da za *vsako* naravno število $a \ge 1$ velja $0^a = 0 \cdot 0^{a-1}$. Podobno je $0! = 1$, saj želimo veljavnost enakosti $a! = a \cdot (a-1)!$ zagotoviti za vsako naravno število $a \ge 1$.

zadnjo enakost utemeljimo z uporabo (6.12).

S tem pa je dokazan tudi izrek. Prispevek elementa a v desni strani enakosti (6.10) je enak 1 natanko tedaj, ko a pripada uniji na levi strani. Sicer je enak 0. □

Poglavje 7

Osnove teorije števil

V tem poglavju bomo obravnavali računanje s celimi števili.

Z računalniškega in računalnikarjevega stališča ima celoštevilska aritmetika veliko prednost pred aritmetiko z realnimi števili (predstavljenimi v premični piki). S celimi števili lahko računamo natančno, popolnoma brez zaokrožitvenih napak. Še vedno sicer obstaja nevarnost prekoračitve obsega, ki pa jo po eni strani precej lažje ukrotimo, po drugi strani pa imamo z njo opravka tudi v premični piki.

Seveda so s celoštevilsko aritmetiko povezane omejitve. Nekateri problemi, denimo modeliranje fizikalnih sistemov, težko shajajo brez realnih števil (kakorkoli jih že predstavimo). Če se bomo omejili na računanje v množici celih števil $\mathbb{Z}$, se bomo implicitno tudi omejili na probleme, ki jih s celoštevilskimi računi zadovoljivo obvladujemo.

7.1 Celi del realnega števila

Na kakšen način lahko realno število preoblikujemo v celo število? Prva ideja je zaokroževanje, realno število a približimo z a najbližjim celim številom. Takšen približek je lahko večji ali manjši od začetnega števila a.

V nekaterih primerih bo ugodno obdržati nadzor nad predznakom razlike. Tako definiramo preslikavi *spodnji* in *zgornji celi del* realnega števila, preslikavi

$$\lfloor . \rfloor : \mathbb{R} \to \mathbb{Z} \quad \text{in} \quad \lceil . \rceil : \mathbb{R} \to \mathbb{Z}, \tag{7.1}$$

definirani z naslednjima predpisoma

$$\begin{aligned} \lfloor a \rfloor &= \max\{n \mid n \in \mathbb{Z} \text{ in } n \leq a\}, \\ \lceil a \rceil &= \min\{n \mid n \in \mathbb{Z} \text{ in } n \geq a\}. \end{aligned}$$

Za realno število a je $\lfloor a \rfloor$ *spodnji celi del števila* a in $\lceil a \rceil$ *zgornji celi del števila* a.

Kot zgled v preslikavi vstavimo števili π in $-\pi$.

$$\lfloor \pi \rfloor = 3 \qquad \lceil \pi \rceil = 4 \qquad \lfloor -\pi \rfloor = -4 \qquad \lceil -\pi \rceil = -3$$

Brez dokaza naštejmo nekaj lastnosti zgornjega in spodnjega celega dela, pri čemer z a označimo poljubno realno število.

(1) $a \in \mathbb{Z} \ \sim \ \lfloor a \rfloor = a \ \sim \ \lceil a \rceil = a$,

(2) $a - 1 < \lfloor a \rfloor \leq a < \lfloor a \rfloor + 1$,

(3) $\lceil a \rceil - 1 < a \leq \lceil a \rceil < a + 1$,

(4) če je $k \in \mathbb{Z}$, potem je $\lfloor a + k \rfloor = \lfloor a \rfloor + k$ in $\lceil a + k \rceil = \lceil a \rceil + k$,

(5) $\lceil a \rceil = -\lfloor -a \rfloor$ in $\lfloor a \rfloor = -\lceil -a \rceil$,

(6) če $a \notin \mathbb{Z}$, potem je $\lceil a \rceil - \lfloor a \rfloor = 1$.

7.2 Deljivost celih števil

Začnimo s klasičnim izrekom o deljivosti celih števil.

Izrek 7.1 (o deljenju) *Naj bosta $n, m \in \mathbb{Z}$ celi števili, pri čemer velja $m \geq 1$. Potem obstajata enolično določeni celi števili $k, r \in \mathbb{Z}$, ki ustrezata enakosti*

$$n = k \cdot m + r, \text{ pri čemer velja } \ 0 \leq r < m. \tag{7.2}$$

Dokaz. Izrek sicer poznamo. Vseeno navedimo kratko ilustracijo, vsaj za primer, ko je $n \geq 0$. V aritmetičnem zaporedju celih števil

$$0 \cdot m, \quad 1 \cdot m, \quad 2 \cdot m, \quad 3 \cdot m, \quad 4 \cdot m, \ \ldots$$

izberimo največje število $k \cdot m$, ki ne presega n. Velja torej $k \cdot m \leq n$ in $(k+1) \cdot m > n$.

Zato za število $r = n - k \cdot m$ velja $r \geq 0$ in $r < m$. Tako določeni celi števili k, r ustrezata (7.2). □

Števili k in r, ki ustrezata (7.2), imenujemo *celoštevilski kvocient* in *ostanek* pri deljenju števila n s številom m.

Kvocient k lahko izračunamo po formuli $k = \lfloor \frac{n}{m} \rfloor$. Ostanek r pa označimo tudi z zapisom

$$r = n \bmod m. \tag{7.3}$$

Relacijo *deljivosti* v množici celih števil definiramo z naslednjim opisom. Za celi števili m in n pravimo, da *m deli n* ali tudi *n je večkratnik m*, kar označimo z $m \mid n$, natanko tedaj, ko je v množici celih števil rešljiva enačba

$$m \cdot x = n. \tag{7.4}$$

Če je $m \geq 1$, potem $m \mid n$ pomeni natanko to, da je ostanek pri deljenju števila n z m enak 0, $n \bmod m = 0$.

Pri ničli je zgodba za odtenek drugačna od kategorične trditve "*z nič deliti ne smemo*".

Vsako celo število m deli 0, oziroma 0 je večkratnik vsakega celega števila m, saj ima pri poljubnem $m \in \mathbb{Z}$ enačba $m \cdot x = 0$ rešitev $x = 0$. Po drugi strani pa število 0 deli samó sámo sebe, saj je enačba $0 \cdot x = n$ rešljiva samo v primeru, ko je $n = 0$.

Izberimo (ne nujno različni) celi števili $m, n \in \mathbb{Z}$. Množico skupnih deliteljev števil m in n označimo z $\mathcal{D}_{m,n}$,

$$\mathcal{D}_{m,n} = \{d \in \mathbb{N} \mid d \mid m \ \text{ in } \ d \mid n \ \text{ in } \ d \geq 0\}.^{1} \tag{7.5}$$

Množica $\mathcal{D}_{m,n}$ je neprazna, saj naravno število 1 deli vsa cela števila. Če je vsaj eno od števil m, n različno od 0, potem je množica $\mathcal{D}_{m,n}$ končna (če je $m \neq 0$, potem so vsi elementi $\mathcal{D}_{m,n}$ navzgor omejeni z $|m|$). *Največji skupni delitelj* števil m in n, $\gcd(m, n)$, je definiran kot

$$\gcd(m, n) = \begin{cases} \max \mathcal{D}_{m,n}, & m \neq 0 \ \text{ ali } \ n \neq 0, \\ 0, & m = n = 0. \end{cases} \tag{7.6}$$

Množico nenegativnih skupnih večkratnikov števil m in n označimo z $\mathcal{V}_{m,n}$,

$$\mathcal{V}_{m,n} = \{v \in \mathbb{Z} \mid v \geq 0 \ \text{ in } \ m \mid v \ \text{ in } \ n \mid v\}. \tag{7.7}$$

Množica $\mathcal{V}_{m,n}$ je neprazna. Pri tem je različna od $\{0\}$ natanko tedaj, ko sta oba, m in n, različna od 0. *Najmanjši skupni večkratnik* števil m in n, $\operatorname{lcm}(m, n)$, definiramo z opisom

$$\operatorname{lcm}(m, n) = \begin{cases} \min \mathcal{V}_{m,n} \setminus \{0\}, & m \neq 0 \ \text{ in } \ n \neq 0, \\ 0, & m = 0 \ \text{ ali } \ n = 0. \end{cases} \tag{7.8}$$

Triznakovni okrajšavi gcd in lcm izvirata iz angleščine.[2]

Deljivost v množici $\mathbb{N}$

Relacija deljivosti v množici celih števil ni delna urejenost. Ker velja, na primer, $5 \mid -5$ in $-5 \mid 5$, relacija deljivosti namreč ni antisimetrična.

V množici naravnih števil $\mathbb{N}$ je relacija deljivosti delna urejenost. Vsa naravna števila delijo ničlo, zato je 0 največji (in posledično tudi edini maksimalni) element v $\mathbb{N}$ glede na

[1] Čeprav, denimo, -5 deli tako 10 kot -15, je množico $\mathcal{D}_{m,n}$ prikladno obdržati znotraj $\mathbb{N}$.
[2] *Greatest common divisor* in *least common multiple*.

relacijo *deljivosti*. Po drugi strani enica 1 deli vsa naravna števila, zato je 1 najmanjši (in edini minimalen) element množice $\mathbb{N}$ glede na relacijo deljivosti.

Za $m, n \in \mathbb{N} \setminus \{0\}$ iz zveze $m \mid n$ sledi $m \leq n$.

Za $m, n \in \mathbb{N}$ sta množici $\mathcal{D}_{m,n}$ oziroma $\mathcal{V}_{m,n}$ družini spodnjih oziroma zgornjih mej za relacijo deljivosti množice $\{m, n\}$. Števili $\gcd(m, n)$ ter $\operatorname{lcm}(m, n)$ pa sta natančni spodnja in zgornja meja množice $\{m, n\}$ glede na relacijo deljivosti.

Kdaj je število n neposredni naslednik števila m glede na deljivost? Kaj so neposredni nasledniki enice?

Zadnje vprašanje ima relativno enostaven odgovor. Če je a sestavljeno[3] število, denimo da a lahko zapišemo kot netrivialen (brez faktorjev enakih 1) produkt $a = bc$, potem a ni neposredni naslednik enice. Velja namreč $1 \mid b$ in $b \mid a$. Vsako praštevilo $p \in \mathbb{P}$ je neposredni naslednik enice glede na relacijo deljivosti. V nasprotnem primeru iz $1 \mid q$ in $q \mid p$, kjer $q \neq 1$ in $q \neq p$, sledi, da p ni praštevilo.

V splošnem velja zveza, da je naravno število n neposredni naslednik števila m (poudarimo: glede na relacijo deljivosti) natanko tedaj, ko je kvocient n/m praštevilo.

Naravno število 0 ni neposredni naslednik (glede na relacijo deljivosti) nobenega naravnega števila. To je nekoliko presenetljivo, saj 0 v relaciji deljivosti nastopa čisto na vrhu strukture (za razliko od standardne urejenosti po velikosti, ko 0 ni neposredni naslednik nobenega naravnega števila iz preprostega razloga, ker je najmanjši element).

Brez dokaza navedimo še nekaj lastnosti operacij gcd in lcm, v naslednjih zvezah a, b, c predstavljajo poljubna cela števila.

(a) gcd in lcm sta *idempotentni operaciji*,

$$\gcd(a, a) = a \quad \text{in} \quad \operatorname{lcm}(a, a) = a,$$

(b) gcd in lcm sta *komutativni operaciji*,

$$\gcd(a, b) = \gcd(b, a) \quad \text{in} \quad \operatorname{lcm}(a, b) = \operatorname{lcm}(b, a),$$

(c) gcd in lcm sta *asociativni operaciji*,

$$\gcd(a, \gcd(b, c)) = \gcd(\gcd(a, b), c) \quad \text{in} \quad \operatorname{lcm}(a, \operatorname{lcm}(b, c)) = \operatorname{lcm}(\operatorname{lcm}(a, b), c),$$

(d) za gcd in lcm veljata oba *distributivnostna zakona*,

$$\gcd(\operatorname{lcm}(a, b), \operatorname{lcm}(a, c)) = \operatorname{lcm}(a, \gcd(b, c)) \quad \text{in}$$
$$\operatorname{lcm}(\gcd(a, b), \gcd(a, c)) = \gcd(a, \operatorname{lcm}(b, c)).$$

[3]Pojem *sestavljenega števila* in *praštevila* bomo natančneje spoznali v enem od prihodnjih razdelkov.

Zaradi asociativnosti lahko operaciji gcd in lcm smiselno razširimo na večje število parametrov. Za, na primer, štiri cela števila a, b, c, d lahko definiramo $\gcd(a, b, c, d)$ in $\operatorname{lcm}(a, b, c, d)$ s predpisoma

$$\gcd(a, b, c, d) = \gcd(a, \gcd(b, \gcd(c, d))), \tag{7.9}$$
$$\operatorname{lcm}(a, b, c, d) = \operatorname{lcm}(a, \operatorname{lcm}(b, \operatorname{lcm}(c, d))). \tag{7.10}$$

Kot zgled poiščimo število števil na celoštevilskem intervalu $\{1, \ldots, 1000\}$, ki so deljiva z natančno tremi od števil $3, 5, 7$ in 9.

Ključ do rešitve takšne naloge je v izbiri kvalitetnih oznak. Z A_n označimo množico števil s celoštevilskega intervala, ki so deljiva z n,

$$A_n = \{k \mid 1 \leq k \leq 1000 \text{ in } n \mid k\}.$$

Iščemo moč množice

$$\begin{aligned} A = ((A_3 \cap A_5 \cap A_7) \setminus A_9) \cup ((A_5 \cap A_7 \cap A_9) \setminus A_3) \cup \\ \cup ((A_7 \cap A_9 \cap A_3) \setminus A_5) \cup ((A_9 \cap A_3 \cap A_5) \setminus A_7) \end{aligned} \tag{7.11}$$

Členi unije (7.11), množice

$$(A_3 \cap A_5 \cap A_7) \setminus A_9, \ (A_5 \cap A_7 \cap A_9) \setminus A_3, \ (A_7 \cap A_9 \cap A_3) \setminus A_5 \quad \text{in} \quad (A_9 \cap A_3 \cap A_5) \setminus A_7,$$

so paroma disjunktni, saj je, denimo, množica $(A_7 \cap A_9 \cap A_3) \setminus A_5$ disjunktna z A_5, medtem ko je množica $(A_9 \cap A_3 \cap A_5) \setminus A_7$ vsebovana v A_5. Opazimo tudi vsebovanost $A_9 \subseteq A_3$. Zato je, med drugim, množica $(A_5 \cap A_7 \cap A_9) \setminus A_3$ prazna, saj je $A_5 \cap A_7 \cap A_9 \subseteq A_9 \subseteq A_3$.

Zato je

$$|A| = |(A_3 \cap A_5 \cap A_7) \setminus A_9| + |(A_7 \cap A_9) \setminus A_5| + |(A_9 \cap A_5) \setminus A_7|$$

V preseku množic $A_m \cap A_n$ so natančno tista števila, ki so deljiva tako z m kot tudi z n. Zato je $A_m \cap A_n = A_{\operatorname{lcm}(m,n)}$. Torej lahko račun nadaljujemo takole.

$$\begin{aligned} |A| &= |A_{\operatorname{lcm}(3,5,7)} \setminus A_9| + |A_{\operatorname{lcm}(7,9)} \setminus A_5| + |A_{\operatorname{lcm}(5,9)} \setminus A_7| \\ &= |A_{105} \setminus A_9| + |A_{63} \setminus A_5| + |A_{45} \setminus A_7| \\ &= |A_{105}| - |A_{105} \cap A_9| + |A_{63}| - |A_{63} \cap A_5| + |A_{45}| - |A_{45} \cap A_7| \\ &= |A_{105}| - |A_{\operatorname{lcm}(105,9)}| + |A_{63}| - |A_{\operatorname{lcm}(63,5)}| + |A_{45}| - |A_{\operatorname{lcm}(45,7)}| \\ &= |A_{105}| - |A_{315}| + |A_{63}| - |A_{315}| + |A_{45}| - |A_{315}| \end{aligned}$$

Pri tem smo upoštevali, da je

$$\operatorname{lcm}(3, 5, 7) = 105, \quad \operatorname{lcm}(7, 9) = 63, \quad \operatorname{lcm}(5, 9) = 45 \quad \text{in} \quad \operatorname{lcm}(3, 5, 7, 9) = 315.$$

Za moč množice A_n pa lahko uporabimo formulo

$$|A_n| = \left\lfloor \frac{1000}{n} \right\rfloor ,$$

saj je na celoštevilskem intervalu $\{1, \ldots, 1000\}$ natančno $\lfloor 1000/n \rfloor$ večkratnikov števila n. Zato je

$$\begin{aligned} |A| &= \left\lfloor \frac{1000}{105} \right\rfloor - \left\lfloor \frac{1000}{315} \right\rfloor + \left\lfloor \frac{1000}{63} \right\rfloor - \left\lfloor \frac{1000}{315} \right\rfloor + \left\lfloor \frac{1000}{45} \right\rfloor - \left\lfloor \frac{1000}{315} \right\rfloor \\ &= 9 - 3 + 15 - 3 + 22 - 3 = 37. \end{aligned}$$

Končno lahko zapišemo odgovor.

> Na celoštevilskem intervalu $\{1, \ldots, 1000\}$ obstaja natanko 37 števil, ki so deljiva z natančno tremi od števil $3, 5, 7, 9$.

Razširjeni Evklidov algoritem

Izračun največjega skupnega delitelja bomo zaupali *razširjenemu Evklidovemu algoritmu REA*.

Kar na začetku zožimo naš zorni kot. Če je $m < 0$, potem je $\gcd(m, n) = \gcd(|m|, n)$, če pa je $m = 0$, potem je $\gcd(m, n) = \gcd(0, n) = |n|$. V primeru enakih pozitivnih argumentov je $\gcd(m, m) = m$, zato se bomo v nadaljni obravnavi (celem preostanku razdelka) omejili na primer, ko je

$$m > 0\,, \quad n > 0 \quad \text{in} \quad m > n. \tag{7.12}$$

Naj bosta m in n naravni števili. *Celoštevilska linearna kombinacija* števil m in n je izraz

$$s \cdot m + t \cdot n, \tag{7.13}$$

kjer sta s in t celi števili. Imenujemo ju tudi *koeficienta* celoštevilske linearne kombinacije (7.13). Tudi za vrednost izraza $s \cdot m + t \cdot n$ bomo uporabljali isti tèrmin *celoštevilska linearna kombinacija* števil m in n.

Z uporabo razširjenega Evklidovega algoritma bomo izračunali največji skupni delitelj števil 765 in 646. Postopek bo potekal okoli ovinka. Raje kot iskanje deliteljev števil 765 in 646 bomo zapisovali celoštevilske linearne kombinacije števil 765 in 646, ki imajo strogo pozitivne, a kar se da majhne, vrednosti.

Zapis algoritma bomo organizirali po vrsticah. V i-ti vrstici bo zapisana enačba-celoštevilska linearna kombinacija števil 765 in 636 oblike

$$s_i \cdot 765 + t_i \cdot 646 = r_i.$$

Vrstice-enačbe označimo z zaporednimi rimskimi številkami, vrednosti vrstic r_i pa imenujemo ostanki.

Prvi dve vrstici-enačbi I in II sta predpisani. V prvi vrstici I s koeficientoma 1 in 0 izrazimo število 765, v drugi vrstici II s koeficientom 0 in 1 izrazimo 646. Števili 765 in 646 ravno tako imenujemo ostanka, $r_1 = 765$, $r_2 = 646$.

Kako iz dveh zaporednih vrstic dobimo naslednjo? Induktivno privzemimo, da smo zapisali vrstici-enačbi, ki izražata ostanka r_{i-1} in r_i, pri čemer naj velja $r_{i-1} > r_i > 0$. Naslednji ostanek r_{i+1} izračunamo kot ostanek pri deljenju r_{i-1} z r_i,

$$r_{i+1} = r_{i-1} \bmod r_i, \tag{7.14}$$

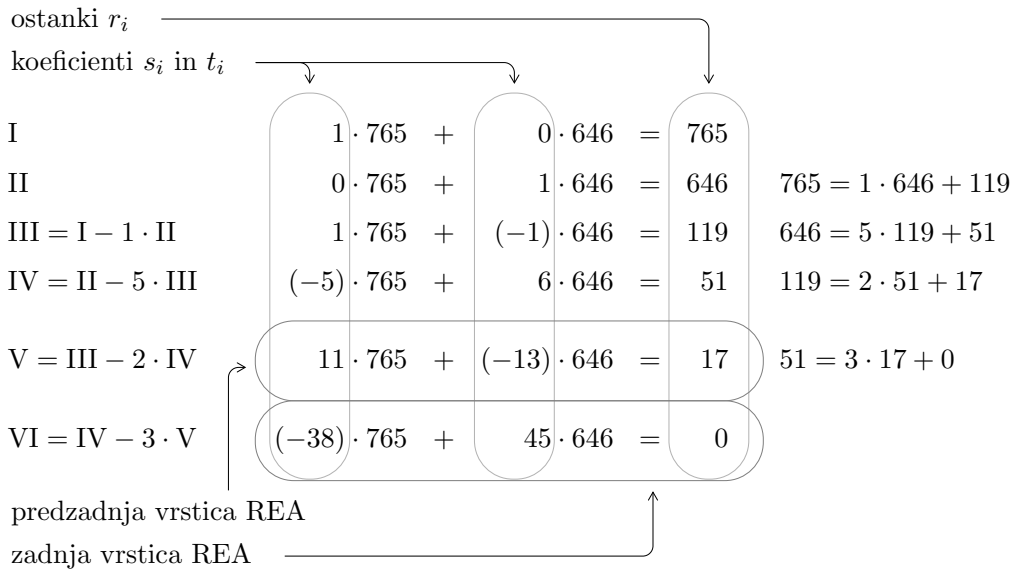
oziroma

$$r_{i+1} = r_{i-1} - k \cdot r_i, \tag{7.15}$$

kjer je k celoštevilski kvocient $\lfloor r_{i-1}/r_i \rfloor$. Iz definicije ostanka pri deljenju sledi, da je $r_{i+1} < r_i$.

Kdaj induktivno zapisanega postopka ni moč nadaljevati? V primeru, ko je najnovejši izraženi ostanek enak 0. Indeks takšne vrstice označimo s črko z. Torej je zadnja vrstica REA tista, v kateri je ostanek $r_z = 0$, predzadnja vrstica pa je vrstica z *zadnjim neničelnim* ostankom r_{z-1}.

Zdaj pa le zapišimo račune. Na levi strani postopka opišemo, kako trenutno vrstico-enačbo izrazimo iz dveh vrstic, ki ležita neposredno nad trenutno vrstico. Na desni strani so računi celoštevilskih deljenj po dveh zaporednih ostankov, ki določijo celoštevilske kvociente k_i. Iste kvociente uporabimo za izračun naslednje vrstice-enačbe.



Za predstavljeni zgled REA veljajo naslednje lastnosti.

(1) *Vsak ostanek r_i, $i = 1, \ldots, z$, je celoštevilska linearna kombinacija števil $m = 765$ in $n = 646$.*

Vsaka posamezna vrstica namreč izrazi ostanek r_i na opisani način.

(2) *Zadnji neničelni ostanek $r_{z-1} = 17$ deli vse ostanke $r_z, r_{z-1}, \ldots, r_2, r_1$. Posebej, zadnji neničelni ostanek je skupni delitelj števil $m = r_1 = 765$ in $n = r_2 = 646$.*

Dokazujemo induktivno, najprej bomo utemeljili naslednjo implikacijo. Če zadnji neničelni ostanek $r_{z-1} = 17$ deli dva zaporedna ostanka r_{i+1} in r_i, potem r_{z-1} deli tudi ostanka r_i in r_{i-1}.

Večjih težav ni. Če r_{z-1} deli r_{i+1} in r_i, potem mora deliti tudi vsako njuno celoštevilsko linearno kombinacijo. Ostanek r_{i-1} pa lahko zaradi zveze (7.15) izrazimo kot celoštevilsko linearno kombinacijo ostankov r_i in r_{i+1}.

Končno, ker $r_{z-1} = 17$ deli tako $r_z = 0$ in $r_{z-1} = 17$, lahko induktivno pridelamo, da r_{z-1} deli vse ostanke $r_z, r_{z-1}, \ldots, r_2, r_1$.

In še, tole gre skozi brez dokaza.

(3) *Če je d skupni delitelj števil $m = r_1 = 765$ in $n = r_2 = 646$, potem d deli tudi vsako njuno celoštevilsko linearno kombinacijo. Posebej, d deli zadnji neničelni ostanek $r_{z-1} = 17$.*

Zgornja analiza je sicer zapisana s konkretnimi števili 765, 646 in 17. Če jih iz (1), (2) in (3) izbrišemo, dobimo dokaz, ki velja tudi za splošna m in n. Vse skupaj lahko strnemo v en sam krovni izrek REA.

Izrek 7.2 (REA) *Naj bosta m in n celi števili. Potem je $\gcd(m, n)$ enak zadnjemu neničelnemu ostanku r_{z-1} REA, ki $\gcd(m, n)$ zapiše tudi kot celoštevilsko linearno kombinacijo m in n,*

$$s_{z-1}m + t_{z-1}n = \gcd(m, n).$$

Zato je

$$\gcd(765, 646) = 11 \cdot 765 + (-13) \cdot 646 = 17.$$

Omenimo še, brez dokaza, da lahko najmanjši skupni večkratnik *števil* m in n preberemo iz *zadnje vrstice* REA. Najmanjši skupni večkratnik števil 765 in 464 je torej enak

$$\operatorname{lcm}(765, 646) = 38 \cdot 765 = 45 \cdot 646 = 29070.$$

Tuja števila

Celi števili m in n sta *(paroma) tuji*, če je $\gcd(m, n) = 1$. V tem primeru pišemo $m \perp n$.

Število 1 je tuje vsem celim številom, tudi samemu sebi. Pokažimo, kot zgled, da sta števili 767 in 646 tuji, izračunajmo njun največji skupni delitelj. Račun nam lahko služi za dodatno vajo iz REA.

I	$1 \cdot 767$	$+$	$0 \cdot 646$	$=$	767	
II	$0 \cdot 767$	$+$	$1 \cdot 646$	$=$	646	$767 = 1 \cdot 646 + 121$
III $=$ I $- 1 \cdot$ II	$1 \cdot 767$	$+$	$(-1) \cdot 646$	$=$	121	$646 = 5 \cdot 121 + 41$
IV $=$ II $- 5 \cdot$ III	$(-5) \cdot 767$	$+$	$6 \cdot 646$	$=$	41	$121 = 2 \cdot 41 + 39$
V $=$ III $- 2 \cdot$ IV	$11 \cdot 767$	$+$	$(-13) \cdot 646$	$=$	39	$41 = 1 \cdot 39 + 2$
VI $=$ IV $- 1 \cdot$ V	$(-16) \cdot 767$	$+$	$19 \cdot 646$	$=$	2	$39 = 19 \cdot 2 + 1$
VII $=$ V $- 19 \cdot$ VI	$315 \cdot 767$	$+$	$(-374) \cdot 646$	$=$	1	$2 = 2 \cdot 1 + 0$
VIII $=$ VI $- 2 \cdot$ VII	$(-646) \cdot 767$	$+$	$767 \cdot 646$	$=$	0	

Nadaljujmo s tehnično lemo.

Lema 7.3 *Naj bosta m in n celi števili in $m > 0$. Potem velja:*

(i) $\gcd(m, n) = \gcd(m, n \bmod m)$,

(ii) $m \perp n$ *natanko tedaj, ko je* $m \perp (n \bmod m)$

Dokaz. Za (i) upoštevamo, da je $n \bmod m = n - k \cdot m$, kjer je k celoštevilski kvocient števil n in m. Število $n \bmod m$ je torej celoštevilska linearna kombinacija števil m in n. Velja pa tudi obrat, tudi n lahko zapišemo kot celoštevilsko linearno kombinacijo števil m in $n \bmod m$, saj je

$$n = 1 \cdot (n \bmod m) + k \cdot m.$$

Zato so skupni delitelji števil m in n natančno isti kot skupni delitelji števil m in $n \bmod m$. Posledično velja (i).

Enakovrednost (ii) je enostavna posledica (i). □

Naslednjo trditev tudi že poznamo: če število deli produkt in je tuje enemu od faktorjev, potem mora deliti drugi faktor. Na tem mestu jo lahko celo dokažemo.

Trditev 7.4 *Naj bodo a, b, c cela števila. Če $a \mid b \cdot c$ in je $a \perp b$, potem $a \mid c$.*

Dokaz. Če $a \mid b \cdot c$, potem obstaja $k \in \mathbb{Z}$, pri katerem velja $a \cdot k = b \cdot c$. Če sta števili a in b tuji, potem obstaja njuna celoštevilska linearna kombinacija, ki izrazi enico. Obstajata torej celoštevilska koeficienta $s, t \in \mathbb{Z}$ za katera je $s \cdot a + t \cdot b = 1$.

Računajmo takole:

$$\begin{aligned} c = c \cdot 1 &= c \cdot (s \cdot a + t \cdot b) \\ &= a \cdot s \cdot c + b \cdot c \cdot t \\ &= a \cdot s \cdot c + a \cdot k \cdot t \\ &= a \cdot (s \cdot c + k \cdot t) \end{aligned}$$

Torej število a res deli c. □

Tudi zveza med največjim skupnim deliteljem in najmanjšim skupnim večkratnikom ni od muh.

Izrek 7.5 *Naj bosta $a, b \in \mathbb{N}$. Potem je*

$$\gcd(a, b) \cdot \operatorname{lcm}(a, b) = a \cdot b. \tag{7.16}$$

Dokaz. Če je katero od števil a, b enako 0, potem je $\operatorname{lcm}(a, b) = 0$ in enakost velja. V nadaljevanju privzamemo, da je $a > 0$ in $b > 0$.

Naj bo $d = \gcd(a, b)$. Števili a in b lahko prepišemo kot produkta

$$a = a_1 \cdot d \quad \text{in} \quad b = b_1 \cdot d,$$

pri čemer sta števili a_1 in b_1 paroma tuji.

Pokazati je potrebno, da je število

$$v = \frac{a \cdot b}{d} = a \cdot b_1 = a_1 \cdot b$$

najmanjši skupni večkratnik števil a in b. Iz zgornjih enakosti sledi, da je v skupni večkratnik a in b.

Denimo, da je $\bar{v}$ skupni večkratnik a in b. V tem primeru obstajata celi števili p in q, za kateri je

$$\bar{v} = p \cdot a = p \cdot a_1 \cdot d = q \cdot b = q \cdot b_1 \cdot d.$$

Število a_1 torej deli izraz $q \cdot b_1$. Ker je $a_1 \perp b_1$, po trditvi 7.4 sledi, da a_1 deli q. Faktor q lahko torej prepišemo kot $q = q_1 \cdot a_1$.

Zato je

$$\bar{v} = q \cdot b = q_1 \cdot a_1 \cdot b = q_1 \cdot v,$$

z drugimi besedami: $\bar{v}$ je večkratnik števila v.

Pokazali smo, da je vsak skupni večkratnik števil a in b tudi večkratnik števila v. To pomeni, da je $v = \operatorname{lcm}(a, b)$ in dokaz izreka je pod streho. □

Dolžni smo še dva komentarja. Edini razlog, da zveza (7.16) ne velja za poljubni celi števili a, b, leži v predznaku. Produkt števil $a \cdot b$ je morda negativno število, ki pa je po *absolutni vrednosti* vedno enako produktu $\gcd(a, b) \cdot \operatorname{lcm}(a, b)$.

Formule (7.16) pa tudi v primeru pozitivnih števil a, b, c ne moremo posplošiti na več faktorjev.

$$10 \cdot 12 \cdot 14 = 1680 \neq \gcd(10, 12, 14) \cdot \operatorname{lcm}(10, 12, 14) = 2 \cdot 420 = 840.$$

7.3 Linearne diofantske enačbe

Enačba premice v ravnini se glasi $ax + by = c$, pri čemer so a, b in c realna števila (pa še a in b ne smeta oba hkrati biti enaka 0).

Kaj pa, če nas zanimajo celoštevilske točke na premici, takšne, ki imajo dve celoštevilski koordinati? Če bodo parametri a, b, c čisto poljubni, bo morda celoštevilskih točk precej malo. Sistematično jih lahko pričakujemo v primeru, ko bodo parametri a, b in c celoštevilski.

Linearna diofantska enačba z dvema neznankama (tudi *LDE*) je enačba oblike

$$ax + by = c, \tag{7.17}$$

pri čemer so parametri $a, b, c \in \mathbb{Z}$, ravno tako pa iščemo rešitve v množici celih števil. *Rešitev* enačbe (7.17) je par celih števil x, y, ki izpolni enakost.

Parametra a in b imenujemo tudi *koeficienta*, parametru c pa standardno pravimo *desna stran* enačbe (7.17).

Kot zgled obravnavajmo enačbo

$$21x + 9y = 138. \tag{7.18}$$

Celoštevilski koordinati (morebitne) rešitve x in y imata vlogi koeficientov celoštevilske linearne kombinacije števil 21 in 9. Poskusimo z uporabo REA.

Zakaj? Nas morda zanima največji skupni delitelj števil 21 in 9? V resnici ne, toda z uporabo REA bomo lahko izrazili število 138 kot celoštevilsko linearno kombinacijo koeficientov 21 in 9 (če je to sploh možno).

$$\begin{array}{l|rcrcl|l}
\text{I} & 1 \cdot 21 & + & 0 \cdot 9 & = & 21 & \\
\text{II} & 0 \cdot 21 & + & 1 \cdot 9 & = & 9 & 21 = 2 \cdot 9 + 3 \\
\text{III} = \text{I} - 2 \cdot \text{II} & 1 \cdot 21 & + & (-2) \cdot 9 & = & 3 & 9 = 3 \cdot 3 + 0 \\
\text{IV} = \text{II} - 3 \cdot \text{III} & (-3) \cdot 21 & + & 7 \cdot 9 & = & 0 &
\end{array} \tag{7.19}$$

Največji skupni delitelj $\gcd(21, 9) = 3$ smo izrazili kot celoštevilsko linearno kombinacijo števil 21 in 9. Če to linearno kombinacijo (predzadnjo vrstico REA) pomnožimo s

kvocientom $\frac{138}{3} = 46$, pridelamo zvezo

$$46 \cdot 21 + (-92) \cdot 9 = 138,$$

kar pomeni, da je $x_0 = 46, y_0 = -92$ rešitev enačbe (7.18).

Oglejmo si še enačbo z alternativno desno stranjo

$$21x + 9y = 47. \tag{7.20}$$

Zgornji trik nas pripelje v težave. Zdi se, da je potrebno predzadnjo vrstico REA (7.19) pomnožiti s številom $\frac{47}{3}$, da pridelamo par x, y, ki zadosti enačbi (7.20). Toda tako dobljeni števili x, y nista celi!

Težava je sistemske narave. Vsaka celoštevilska linearna kombinacija števil 21 in 9 je po vrednosti večkratnik števila $\gcd(21, 9) = 3$. Desna stran enačbe (7.20) pa ni deljiva s 3, zato enačba (7.20) ni rešljiva.

Postopek reševanja, ki smo ga uporabili na primerih LDE (7.18) in (7.20), v prvo uspešno in potem še neuspešno, lahko ponovimo na poljubni LDE z dvema neznankama. Dokazali smo torej naslednji rezultat.

Izrek 7.6 *Linearna diofantska enačba*

$$ax + by = c$$

je rešljiva natanko tedaj, ko največji skupni delitelj koeficientov $\gcd(a, b)$ *deli desno stran* c.

Izrek 7.6 uspešno loči rešljive LDE z dvema neznankama od nerešljivih. Še vedno pa bi želeli poiskati *vse* rešitve enačbe (7.17), če vemo, da kakšna rešitev obstaja.

Denimo, da para x_1, y_1 in x_2, y_2 rešita enačbo (7.17). Privzemimo, iz čisto tehničnih razlogov, da sta koeficienta $a, b \in \mathbb{N}$, morebiten drugačen predznak lahko vedno skrijemo h koordinati rešitve. Potem je

$$\begin{aligned} 0 = c - c &= (ax_1 + by_1) - (ax_2 + by_2) \\ &= a(x_1 - x_2) + b(y_1 - y_2). \end{aligned}$$

S preurejanjem pridelamo enakost

$$a(x_1 - x_2) = b(y_2 - y_1).$$

Izraz $a(x_1 - x_2)$ je po eni strani večkratnik števila a, po drugi strani pa večkratnik števila b. Zato je večkratnik števila $\operatorname{lcm}(a, b)$ in z upoštevanjem izreka 7.5 lahko zapišemo

$$a(x_1 - x_2) = b(y_2 - y_1) = t \cdot \operatorname{lcm}(a, b) = t \cdot \frac{ab}{\gcd(a, b)} \tag{7.21}$$

Dokazali smo naslednji izrek.

Izrek 7.7 *Denimo, da je linearna diofantska enačba*

$$ax + by = c$$

rešljiva in je par x_0, y_0 njena rešitev. Vse rešitve enačbe dobimo s formulo

$$x_t = x_0 + t \cdot \frac{b}{\gcd(a,b)},$$
$$y_t = y_0 - t \cdot \frac{a}{\gcd(a,b)},$$

kjer je t poljuben celoštevilski parameter.

Kako vse rešitve LDE poiščemo v praksi? Katere so vse rešitve LDE (7.18)?

Eno rešitev smo poiskali tako, da smo predzadnjo vrstico REA pomnožili s kvocientom $\frac{138}{3} = 46$. Če tako dobljeni celoštevilski linearni kombinaciji prištejemo še poljuben večkratnik *zadnje vrstice*, desne strani ne pokvarimo. Koordinati rešitve pa lahko neposredno preberemo.

$$\begin{aligned} 138 &= \frac{138}{3} \cdot 3 + t \cdot 0 \\ &= 46 \cdot (1 \cdot 21 + (-2) \cdot 9) + t \cdot ((-3) \cdot 21 + 7 \cdot 9) \\ &= 46 \cdot 21 + (-92) \cdot 9 + (-3t) \cdot 21 + 7t \cdot 9 \\ &= (46 - 3t) \cdot 21 + (-92 + 7t) \cdot 9 \end{aligned}$$

Odtod preberemo množico *vseh* rešitev LDE (7.18):

$$x_t = 46 - 3t,$$
$$y_t = -92 + 7t,$$

pri čemer je t poljuben celoštevilski parameter.

Za konec brez dokaza navedimo še pogoj za rešljivost linearne diofantske enačbe z večjim številom neznank. Identičen je tistemu iz izreka 7.6.

Izrek 7.8 *Linearna diofantska enačba*

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots a_kx_k = c$$

je rešljiva natanko tedaj, ko

$$\gcd(a_1, a_2, a_3, \ldots, a_k) \quad \textit{deli desno stran } c.$$

7.4 Praštevila

Naravno število n je *praštevilo*, če ima natančno *dva* delitelja v množici naravnih števil. Naravno število $n > 0$ je *sestavljeno število*, če ima vsaj tri delitelje v množici naravnih števil.

Vsako naravno število n je deljivo s samim sabo, zato ima vsak $n \in \mathbb{N}$ vsaj enega delitelja v $\mathbb{N}$. Enica 1 je edino naravno število, ki ima v množici $\mathbb{N}$ samo enega delitelja. Ničla 0 je ravno tako nekaj posebnega, deljiva je z vsakim naravnim številom, pa je vseeno ne tlačimo v družbo sestavljenih števil.

Preostala naravna števila iz $\mathbb{N} \setminus \{0, 1\}$ pa spadajo v natanko enega od dveh predalčkov. Vsako naravno število $n \geq 2$ je bodisi praštevilo bodisi sestavljeno število.

Praštevila, ki so manjša od 100, so natančno

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89 \text{ in } 97.$$

Število 2 je edino sodo praštevilo, preostala praštevila so liha.

Množico praštevil bomo označili s $\mathbb{P}$.

Trditev 7.9 *Naj bo p praštevilo in a, b celi števili.*

(i) *Bodisi je $p \perp a$ bodisi $p \mid a$.*

(ii) *Če $p \mid (a \cdot b)$, potem $p \mid a$ ali $p \mid b$.*

(iii) *Če p deli celoštevilski produkt $a_1 \cdot a_2 \cdot \cdots \cdot a_k$, potem p deli vsaj eno od števil $a_1, \ldots, a_k$.*

Dokaz. Števili 1 in p sta edina pozitivna delitelja praštevila p. Zato je največji *skupni* delitelj $\gcd(a, p)$ enak bodisi 1 bodisi p. V prvem primeru je $p \perp a$, v drugem primeru pa $p \mid a$.

Za dokaz (ii) lahko privzamemo, da p ne deli faktorja a. Po (i) sledi $p \perp a$. Zdaj pa z uporabo trditve 7.4 sledi, da $p \mid b$.

Pri (iii) sklepamo induktivno, praštevilo p deli vsaj enega od faktorjev $(a_1 \cdot \cdots \cdot a_{k-1})$ ali a_k. □

Lema 7.10 *Vsako naravno število $n \geq 2$ je deljivo s katerim od praštevil $p \in \mathbb{P}$.*

Dokaz. Dokazujemo z indukcijo. Števila 2, 3 in 4 so vsa deljiva s katerim od praštevil. Ti majhni zgledi nam služijo kot baza indukcije.

Izberimo naravno število $n \geq 5$ in privzemimo, da za vsako naravno število n', ki zadošča $n' < n$ in $n' \geq 2$, obstaja praštevilo p', ki deli n'.

Če je $n \in \mathbb{P}$, potem je število n deljivo s praštevilom n. Če pa n ni praštevilo, ga lahko zapišemo kot produkt $n = n_1 \cdot n_2$, v katerem sta oba faktorja netrivialna, $n_1 \geq 2$ in $n_2 \geq 2$. To pomeni, da velja $n_1 < n$ in $n_2 < n$. Po indukcijski predpostavki obstaja praštevilo p_1, ki deli n_1.

Zato je tudi število $n = n_1 \cdot n_2$ deljivo s praštevilom p_1 in lema je dokazana. □

Praštevila so bila v celotni zgodovini matematike izvir zanimivih matematičnih problemov in gonilo matematičnega raziskovanja. Koliko pa je praštevil, se glasi eno od začetnih vprašanj. Odgovor nanj so poznali že v antični Grčiji, Evklidov dokaz izreka pa je še danes zgled elegantnega dokaza z uporabo protislovja.

Izrek 7.11 (Evklid) *Množica praštevil $\mathbb{P}$ je neskončna.*

Dokaz. Privzemimo nasprotno. Denimo, da je množica $\mathbb{P}$ končna. Tedaj lahko zapišemo vsa praštevila v končno zaporedje $p_1, p_2, p_3, \ldots, p_z$. Sestavimo število

$$N = (p_1 \cdot p_2 \cdot p_3 \cdot \cdots \cdot p_z) + 1.$$

Število N ni deljivo z nobenim od praštevil $p_1, \ldots, p_z$, saj pri deljenju števila N s katerimkoli od praštevil $p_1, \ldots, p_z$ dobimo ostanek 1. Po lemi 7.10 pa je število N deljivo z vsaj enim od praštevil iz $\mathbb{P}$. To je protislovje, zato je množica $\mathbb{P}$ neskončna. □

Praštevila predstavljajo osnovne gradnike za gradnjo naravnih števil. Vsako naravno število (razen dveh zelo majhnih 0 in 1) lahko zapišemo kot produkt praštevil (pri čemer dopušC"amo tudi produkte z enim samim faktorjem) in še več.

Izrek 7.12 *Vsako naravno število $n \geq 2$ lahko zapišemo kot produkt praštevil. Če se ne oziramo na vrstni red faktorjev, je ta zapis enoličen.*

Dokaz. Pokažimo najprej obstoj praštevilskega razcepa. Za majhna naravna števila velja

$$2 = 2\,, \quad 3 = 3\,, \quad 4 = 2 \cdot 2\,, \quad 5 = 5\,, \quad 6 = 2 \cdot 3.$$

Pri tem zapis $2 = 2$ pomeni, da smo praštevilo 2 zapisali kot produkt, v katerem je en sam faktor (namreč praštevilo 2).

Dokazujemo z indukcijo. Izberimo naravno število $n \geq 7$ in privzemimo, da lahko vsako naravno število n', ki zadošča $2 \leq n' < n$, zapišemo kot produkt praštevil.

Če je n praštevilo, potem smo z njegovim zapisom zaključili. Pišemo lahko $n = n$. Sicer je n sestavljeno število in po lemi 7.10 obstaja praštevilo $p \in \mathbb{P}$, za katerega smemo pisati $n = pn_1$. Število n_1 je strogo manjše od n (in vsaj 2), zato ga po indukcijski

predpostavki lahko zapišemo kot produkt praštevil. Če v ta produkt dodamo še faktor p, pridelamo zapis naravnega števila n kot produkt praštevil.

Dokaz enoličnosti znova izdelamo z indukcijo. Baza indukcije trdi, da lahko število 2 na en sam način zapišemo kot produkt praštevil. To je res, saj je produkt katerikolih *dveh* naravnih števil ≥ 2 strogo večji od 3. Edini zapis kot praštevilski produkt je že omenjeni $2 = 2$.

Induktivni dokaz izdelajmo do konca na nekoliko drugačen način, z minimalnim protiprimerom. V našem primeru je *minimalni protiprimer* najmanjše naravno število, za katerega besedilo izreka ne velja. Naš cilj je pokazati, da minimalni protiprimer sploh ne obstaja.

Denimo, da je n najmanjše naravno število, ki ga lahko zapišemo kot produkt praštevil na dva (bistveno, ne samo v vrstnem redu) različna načina. Zapišimo dotična praštevilska razcepa,

$$n = p_1 \cdot p_2 \cdot p_3 \cdot \cdots \cdot p_k,$$
$$n = q_1 \cdot q_2 \cdot q_3 \cdot \cdots \cdot q_\ell.$$

Pri tem so števila $p_1, \ldots, p_k, q_1, \ldots, q_\ell$ praštevila. Ker je n minimalni protiprimer, se praštevila v enem od zapisov nikakor ne pojavijo v drugem. Po domače, ne obstajata indeksa i in j, pri katerih bi veljalo $p_i = q_j$. Zakaj?

V nasprotnem primeru bi število $n/p_i = n/q_j$ dopuščalo dva bistveno različna zapisa kot produkt praštevil, kar je v protislovju z minimalnostjo števila n.

Praštevilo p_1 deli produkt $n = q_1 \cdot q_2 \cdot q_3 \cdot \cdots \cdot q_\ell$. Po trditvi 7.9 praštevilo p_1 deli vsaj enega od faktorjev in je hkrati tuje vsem faktorjem. To je nemogoče, minimalni protiprimer torej ne obstaja. □

7.5 Eulerjeva funkcija

Izberimo naravno število 24 in opazujmo, katera izmed naravnih števil v množici $\{1, 2, \ldots, 24\}$ so tuja številu 24. To so natačno števila

$$1, 5, 7, 11, 13, 17, 19 \quad \text{in} \quad 23.$$

V množici celih števil $\{1, 2, \ldots, 24\}$ obstaja natanko 8 števil, ki so tuja številu 24.

Eulerjeva funkcija φ je preslikava

$$\varphi : \mathbb{N} \to \mathbb{N},$$

definirana z opisom

$$\varphi(n) = |\{k \in \mathbb{N} \mid 1 \leq k \leq n \ \text{ in } \ k \perp n\}|, \tag{7.22}$$

$\varphi(n)$ je število naravnih števil na intervalu med 1 in n, ki so tuja n.

Prvi zgled določi $\varphi(24) = 8$. Oglejmo si nekaj začetnih vrednosti funkcije φ. Pri posameznem argumentu n Eulerjeve funkcije podčrtajmo števila, ki *niso* tuja n.

$$\begin{array}{ll} \varphi(0) = 0 & \\ \varphi(1) = 1 & 1 \\ \varphi(2) = 1 & 1, \underline{2} \\ \varphi(3) = 2 & 1, 2, \underline{3} \\ \varphi(4) = 2 & 1, \underline{2}, 3, \underline{4} \\ \varphi(5) = 4 & 1, 2, 3, 4, \underline{5} \\ \varphi(6) = 2 & 1, \underline{2}, \underline{3}, \underline{4}, 5, \underline{6} \end{array}$$

V razdelku se bomo ukvarjali z računanjem vrednosti Eulerjeve funkcije φ.

Če je p praštevilo, potem so vsa naravna števila na intervalu med 1 in $p-1$ tuja praštevilu p, saj p ne deli nobenega od omenjenih števil. Število p ni tuje samemu sebi. Zato velja naslednja trditev:

Trditev 7.13 *Če je p praštevilo, potem je $\varphi(p) = p - 1$.*

Nadaljujmo s praštevilskimi potencami.

Trditev 7.14 *Če je p praštevilo in $n \geq 1$, potem je $\varphi(p^n) = p^n - p^{n-1}$.*

Dokaz. Privzamemo lahko, da je $n \geq 2$. Primer $n = 1$ je predstavljen v trditvi 7.13.

Edini pozitivni delitelji praštevilske potence p^n so praštevilske potence z manjšimi eksponenti $1, p, p^2, \ldots, p^n$. Če naravno število k *ni tuje* številu p^n, potem je

$$\gcd(p^n, k) \in \{p, p^2, \ldots, p^n\}.$$

Zato je p delitelj števila k.

Koliko števil pa je deljivih s p na celoštevilskem intervalu med 1 in p^n? To so natančno večkratniki praštevila p,

$$1 \cdot p, \quad 2 \cdot p, \quad 3 \cdot p, \ \ldots, \quad p^{n-1} \cdot p,$$

in njihovo število je natančno p^{n-1}. Zato je $\varphi(p^n) = p^n - p^{n-1}$. □

Zadnje orodje govori o produktu paroma tujih števil.

Trditev 7.15 *Če sta neničelni naravni števili a in b tuji, potem je*

$$\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b).$$

Dokaz. Števila med 1 in $a \cdot b$ zapišimo v tabelo z b vrsticami in a stolpci.

$$\begin{array}{rrrcr} 1 & 2 & 3 & \dots & a \\ a+1 & a+2 & a+3 & \dots & 2a \\ 2a+1 & 2a+2 & 2a+3 & \dots & 3a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (b-1)a+1 & (b-1)a+2 & (b-1)a+3 & \dots & ba \end{array} \tag{7.23}$$

Števila, ki so tuja produktu ab, morajo biti tuja a in tudi b. Števila v r-tem stolpcu tabele (7.23) so oblike $ka + r$, pri čemer velja $0 \leq k \leq b-1$. Pri deljenju z a dajo vsa omenjena števila isti ostanek, namreč r. Torej so po Lemi 7.3(ii) bodisi *vsa* tuja a bodisi ni *nobeno* tuje a. Stolpcev, ki vsebujejo številu a tuja števila, je torej natančno $\varphi(a)$.

Pri deljenju z b dajejo števila v r-tem stolpcu same različne ostanke. Iz

$$k_1 a + r \equiv k_2 a + r \pmod{b}$$

namreč sledi $b|(k_1 - k_2) \cdot a$. Po trditvi 7.4 sledi tudi $b|(k_1 - k_2)$, saj je $a \perp b$. Ker pa sta koeficienta k_1, k_2 omejena, velja namreč $0 \leq k_1, k_2 \leq b-1$, je $-b+1 \leq k_1 - k_2 \leq b-1$. Med celimi števili, ki so po absolutni vrednosti strogo manjša od b, je samo ničla deljiva z b. Torej velja $k_1 = k_2$. V vsakem stolpcu je natančno b števil, ki dajejo pri deljenju z b vseh b možnih ostankov.

V tabeli je torej $\varphi(a)$ stolpcev, katerih elementi so tuji a. V vsakem od teh stolpcev pa je $\varphi(b)$ števil, ki so tuja b. Števil, ki so tuja a in tudi b, je v tabeli (7.23) torej natančno $\varphi(a) \cdot \varphi(b)$. □

S pomočjo trditev 7.13, 7.14 in 7.15 lahko izračunamo vrednost Eulerjeve funkcije pri poljubnem naravnem številu $n \geq 1$, če le poznamo njegov praštevilski razcep.

Izrek 7.16 *Naj bo*

$$n = p_1^{k_1} \cdot p_2^{k_2} \cdot \dots \cdot p_m^{k_m}$$

praštevilski razcep števila n, kjer so $p_1, p_2, \dots, p_m$ različna praštevila. Potem je

$$\begin{aligned} \varphi(n) &= (p_1^{k_1} - p_1^{k_1-1}) \cdot (p_2^{k_2} - p_2^{k_2-1}) \cdot \dots \cdot (p_m^{k_m} - p_m^{k_m-1}) \\ &= n \cdot \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdot \dots \cdot \left(1 - \frac{1}{p_m}\right). \end{aligned}$$

Razdelek končajmo z nekaj zgledi izračuna Eulerjeve funkcije φ.

$$\begin{aligned} \varphi(196) &= \varphi(4 \cdot 49) = \varphi(2^2) \cdot \varphi(7^2) = (4-2) \cdot (49-7) = 84 \\ \varphi(720) &= \varphi(16 \cdot 9 \cdot 5) = \varphi(2^4) \cdot \varphi(3^2) \cdot \varphi(5) = 8 \cdot 6 \cdot 4 = 192 \\ \varphi(1200) &= \varphi(16 \cdot 3 \cdot 25) = \varphi(2^4) \cdot \varphi(3) \cdot \varphi(5^2) = 8 \cdot 2 \cdot 20 = 320 \\ \varphi(3125) &= \varphi(5^5) = 5^5 - 5^4 = 3125 - 625 = 2500 \end{aligned}$$

7.6 Modulska aritmetika

V zadnjem razdelku poglavja si bomo ogledali računanje s kongruencami. Izberimo naravno število[4] $m \geq 1$. Celi števili a in b sta *kongruentni po modulu* m, če dasta pri deljenju z m isti ostanek. Pišemo tudi

$$a \equiv b \pmod{m}. \tag{7.24}$$

Po sami definiciji je $a \equiv b \pmod{m}$ enakovredno zvezi $a \bmod m = b \bmod m$, pa tudi dejstvu, da m deli razliko števil a in b, $m \mid (a - b)$.

Denimo, da za celi števili a, b velja

$$a \equiv b \pmod{m}.$$

Potem za vsako celo število c velja

$$\begin{aligned} a \pm c &\equiv b \pm c \pmod{m}, \\ a \cdot c &\equiv b \cdot c \pmod{m}. \end{aligned} \tag{7.25}$$

Če je n naravno število in veljata zvezi

$$a \equiv b \pmod{m} \quad \text{in} \quad c \equiv d \pmod{m},$$

potem veljajo tudi kongruence

$$\begin{aligned} a \pm c &\equiv b \pm d \pmod{m}, \\ a \cdot c &\equiv b \cdot d \pmod{m}, \\ a^n &\equiv b^n \pmod{m}. \end{aligned} \tag{7.26}$$

Kongruenci (7.25) sta posledici kongruenc (7.26), ki ju na hitro utemeljimo. Privzamemo, da m deli izraza $a - b$ in $c - d$. Število m zato deli razliko

$$(a \pm c) - (b \pm d) = (a - b) \pm (c - d).$$

Za razliko produktov se je potrebno znajti s prištevanjem in odštevanjem istega člena $a \cdot d$.

$$a \cdot c - b \cdot d = a \cdot c - a \cdot d + a \cdot d - b \cdot d = a \cdot (c - d) + (a - b) \cdot d$$

Izraz na desni je deljiv z m.

Pravilo krajšanja v splošnem ne velja. Faktorja 6 v kongruenci

$$6 \cdot 5 \equiv 6 \cdot 9 \pmod{8}$$

[4] Izbira $m = 1$ je dopustna, a ne pretirano zanimiva. Smiselno je izbrati večji modul m.

ne moremo krajšati, saj

$$5 \not\equiv 9 \pmod 8.$$

V nekaterih primerih pa smemo vseeno krajšati skupni faktor na obeh straneh kongruence. Če za cela števila a, b, c velja

$$a \cdot c \equiv b \cdot c \pmod m \qquad \text{in je število} \quad c \perp m,$$

potem velja tudi

$$a \equiv b \pmod m. \tag{7.27}$$

Z drugimi besedami, krajšamo lahko faktor, ki je tuj modulu m.

Nadaljujmo z Eulerjevim izrekom. Če je število a tuje modulu m, potem s potenciranjem števila a *sčasoma* pridelamo število, ki je kongruentno 1. Presenetljiv je eksponent, pri katerem se to gotovo zgodi.

Izrek 7.17 (Eulerjev) *Naj bo $a \in \mathbb{Z}$, $m \in \mathbb{N}$, $m \geq 1$ in $a \perp m$. Potem je*

$$a^{\varphi(m)} \equiv 1 \pmod m. \tag{7.28}$$

Dokaz. Trditev očitno velja za $m = 1$, zato smemo privzeti, da je $m \geq 2$. Naj bo

$$\mathcal{T} = \{a_1, a_2, \ldots, a_{\varphi(m)}\}$$

množica vseh števil med 1 in m, ki so tuja modulu m. Definirajmo preslikavo

$$\begin{aligned} &\pi_a : \mathcal{T} \to \mathcal{T}, \\ &\pi_a : x \mapsto a \cdot x \bmod m. \end{aligned}$$

Preverimo najprej, da je π_a dobro definirana. Če $a_i \in \mathcal{T}$, potem je funkcijska vrednost $\pi_a(a_i) = a \cdot a_i \bmod m$ naravno število, ki je strogo manjše od m. Denimo, da je $d = \gcd(a \cdot a_i \bmod m, m) > 1$. V tem primeru število d deli produkt $a \cdot a_i$. Toda števili a in a_i sta obe tuji d (v nasprotnem primeru pridemo v protislovje z $a \perp m$ ali $a_i \perp m$, saj d deli m), kar je v nasprotju s trditvijo 7.4. To pomeni, da je $\gcd(a \cdot a_i \bmod m, m) = 1$ in zato $\pi_a(a_i) = a \cdot a_i \bmod m \in \mathcal{T}$.

Preslikava π_a je tudi injektivna. Če namreč velja

$$\pi_a(a_i) = \pi_a(a_j),$$

potem velja

$$a \cdot a_i \equiv a \cdot a_j \pmod m.$$

Faktor a je tuj modulu m, zato ga lahko v kongruenci krajšamo in pridelamo

$$a_i \equiv a_j \pmod m.$$

To pa je možno samo v primeru enakosti $a_i = a_j$.

Preslikava π_a je torej injektivna in po izreku 6.1 tudi surjektivna in bijektivna. Zato je

$$a_1 \cdot a_2 \cdot a_3 \cdot \cdots \cdot a_{\varphi(m)} \equiv (a \cdot a_1) \cdot (a \cdot a_2) \cdot (a \cdot a_3) \cdot \cdots \cdot (a \cdot a_{\varphi(m)}) \pmod{m}, \quad (7.29)$$

saj je $\pi_a(\mathcal{T}) = \mathcal{T}$.

V enakosti (7.29) lahko krajšamo vse faktorje $a_1, \ldots, a_{\varphi(m)} \in \mathcal{T}$ in pridelamo zaključno enakost

$$1 \equiv a^{\varphi(m)} \pmod{m}.$$

□

Posledica Eulerjevega izreka je mali Fermatov izrek, ki je v resnici več kot 100 let starejšega datuma.

Izrek 7.18 (mali Fermatov) *Naj bo $a \in \mathbb{Z}$ in p praštevilo. Potem je*

$$a^p \equiv a \pmod{p}. \quad (7.30)$$

Dokaz. Če je $a \perp p$, potem po Eulerjevem izreku (izrek 7.17) velja

$$a^p \equiv a \cdot a^{p-1} \equiv a \cdot a^{\varphi(p)} \equiv a \cdot 1 \equiv a \pmod{p}.$$

Če pa $p \mid a$, potem je

$$a^p \equiv 0 \equiv p \pmod{a}.$$

V obeh primerih zveza (7.30) velja. □

Z uporabo lastnosti kongruenc in Eulerjevega izreka lahko izračunamo

$$16^{14^{12^{10}}} \bmod 13. \quad (7.31)$$

Izraz na levi je potenca števila 16. Ker je 16 tuje 13, po izreku 7.17 velja $16^{\varphi(13)} \equiv 16^{12} \equiv 1 \pmod{13}$. Zato pri poljubnih naravnih številih k, r velja zveza

$$16^{12k+r} \equiv \left(16^{12}\right)^k \cdot 16^r \equiv 1^k \cdot 16^r \equiv 16^r \pmod{13}. \quad (7.32)$$

Z drugimi besedami, zanima nas ostanek enostavnejše potence $14^{12^{10}}$ pri deljenju z 12 — torej ostanek po modulu 12.

Eulerjevega izreka ne moremo uporabiti, saj osnova 14 ni tuja modulu 12. Vseeno lahko poskusimo izračunati ostanke nekaj zaporednih potenc števila 14 pri deljenju z 12.

$$\begin{aligned} 14^1 &\equiv 14 \equiv 2 \pmod{12} \\ 14^2 &\equiv 2^2 \equiv 4 \pmod{12} \\ 14^3 &\equiv 2^3 \equiv 8 \pmod{12} \\ 14^4 &\equiv 2^4 \equiv 4 \pmod{12} \end{aligned}$$

Trdimo, da za vsako naravno število $n \geq 2$ velja

$$14^{n+2} \equiv 14^n \pmod{12}. \tag{7.33}$$

Dokazujemo z indukcijo, bazo indukcije smo upravičili z direktnim računom. Indukcijski korak sledi iz množenja zveze (7.33) s 14 na obeh straneh enakosti hkrati.

Odtod sledi, da se ostanki zaporednih potenc števila 14 po modulu 12 ponavljajo s periodo 2.

Potenca 12^{10} je *očitno* večkratnik števila 2, zato pri primerno izbranem k velja $12^{10} = 2k$. Odtod je

$$14^{12^{10}} \equiv 14^{2k} \equiv 14^2 \equiv 4 \pmod{12},$$

z drugimi besedami, število $14^{12^{10}}$ je pri primerno izbranem k' enako $12 \cdot k' + 4$.

Končno je

$$16^{14^{12^{10}}} \equiv 16^{12k'+4} \equiv \left(16^{12}\right)^{k'} \cdot 16^4 \equiv (1^{k'}) \cdot 3^4 \equiv 81 \equiv 3 \pmod{13}$$

7.6.1 Računanje z ostanki

Kongruenčna enačba

$$6 \cdot x \equiv 4 \pmod{8}$$

je rešljiva, saj je število 2 rešitev. Tudi števila oblike $2 + 8k$ so njene rešitve, zato je množica rešitev neskončna.

Vseeno bi želeli natančnejšo informacijo o gostoti rešitev. Je morda ustrezno samo vsako osmo celo število ali so morda rešitve razporejene bolj na gosto?

Izberimo modul m ($m \in \mathbb{N}$ in $m \geq 1$). *Kongruenca po modulu* m je ekvivalenčna relacija v množici celih števil. Kvocientno množico označimo z Z_m,

$$Z_m = \mathbb{Z}/\mathrm{mod}\, m.$$

Ekvivalenčne razrede množice Z_m imenujemo *ostanki po modulu* m. V obravnavi jih bomo enačili z njihovimi predstavniki: posamezen ekvivalenčni razred $[n] \in Z_m$ (množico celih števil, ki dajo pri deljenju z m isti ostanek r) *enačimo* z njegovim predstavnikom $r \in \{0, \ldots, m-1\}$.

Ostanek $a \in Z_m$ je *obrnljiv*, če je rešljiva enačba

$$a \cdot x \equiv 1 \pmod{m}. \tag{7.34}$$

Neničelni ostanek $a \in Z_m$ je *delitelj niča*, če obstaja neničelni ostanek $b \in Z_m$, ki reši enačbo

$$a \cdot x \equiv 0 \pmod{m}. \tag{7.35}$$

Obrnljivi ostanki in delitelji niča porodijo razbitje množice Z_m.

Trditev 7.19 *Vsak ostanek $a \in Z_m$ je bodisi enak 0 bodisi obrnljiv bodisi delitelj niča.*

Dokaz. Definirajmo preslikavo

$$\pi_a : Z_m \to Z_m,$$
$$\pi_a : x \mapsto a \cdot x \quad (= a \cdot x \bmod m).$$

Če je preslikava π_a konstanta 0, potem je $a = 0$. V nasprotnem primeru je $a \neq 0$.

Če je preslikava π_a injektivna (in po izreku 6.1 tudi surjektivna in bijektivna), potem 1 pripada zalogi vrednosti preslikave π_a in je a obrnljiv. V tem primeru a ni delitelj niča, saj je $\pi_a(0) = 0$, preslikava π_a pa je injektivna.

Če preslikava π_a ni injektivna, potem obstajata različna ostanka $b, b' \in Z_m$, za katera je

$$a \cdot b \equiv a \cdot b' \pmod{m}.$$

Privzemimo, da je $b > b'$ in kongruenco prepišemo v

$$a \cdot (b - b') \equiv 0 \pmod{m}.$$

Torej je $a \in Z_m$ delitelj niča. □

Naj bo $a \in Z_m$ obrnljiv ostanek. Če za ostanek $a' \in Z_m$ velja

$$a \cdot a' \equiv 1 \pmod{m},$$

potem je a' *inverz (za množenje)* ostanka a. Inverz ostanka a za množenje bomo označili tudi z a^{-1}.

Iz dokaza trditve 7.19 sledi, da je pri obrnljivem a njegov inverz a^{-1} enolično določen. Ravno tako je v tem primeru a^{-1} obrnljiv in velja $(a^{-1})^{-1} = a$.

Kateri ostanki pa so obrnljivi? Kateri so delitelji niča?

Enačbo (7.34) lahko *prepišemo* v linearno diofantsko enačbo

$$a \cdot x + m \cdot y = 1, \tag{7.36}$$

izraz $a \cdot x$ se od enice 1 razlikuje za večkratnik števila m. Pri tem nas y-koordinata rešitve LDE ne bo zanimala, zanimal nas bo samo ostanek x-koordinate rešitve pri deljenju z m.

Po izreku 7.6 je enačba (7.36) rešljiva natanko tedaj, ko sta števili a in m tuji. To med drugim pomeni, da je v Z_m natančno $\varphi(m)$ obrnljivih ostankov.

Tudi enačbo (7.35) lahko na podoben način prepišemo v linearno diofantsko enačbo

$$a \cdot x + m \cdot y = 0. \tag{7.37}$$

Enačba (7.37) je vedno rešljiva, ostanek a pa je delitelj niča natanko tedaj, ko obstajata rešitvi enačbe (7.37) pri vsaj dveh možnostih x iz celoštevilskega intervala $0, \ldots, m-1$ (0 je vedno rešitev, druga je potem neničelna).

Po izreku 7.7 si x-koordinate zaporednih rešitev diofantske enačbe (7.37) sledijo z razlikami $m/\gcd(a,m)$. Če števili a in m nista tuji, potem je razlika strogo manjša od m. Tako lahko tudi karakteriziramo ločnico med obrnljivimi ostanki in delitelji niča v Z_m.

Trditev 7.20 *Za neničeln ostanek $a \in Z_m \setminus \{0\}$ velja*

(i) *a je obrnljiv natanko tedaj, ko je $\gcd(a,m) = 1$ in*

(ii) *a je delitelj niča natanko tedaj, ko je $\gcd(a,m) \geq 2$.*

Kot zgled izberimo $m = 18$ in določimo obrnljive ostanke in njihove inverze. Število obrnljivih ostankov je enako $\varphi(18) = 6$ in le-ti so

$$1, 5, 7, 11, 13 \quad \text{in} \quad 17.$$

Veljajo naslednje kongruence

$$1 \cdot 1 \equiv 5 \cdot 11 \equiv 7 \cdot 13 \equiv 17 \cdot 17 \equiv 1 \pmod{18}.$$

Po modulu 18 je ostanek 1 inverz samega sebe, ostanka 5 in 11 sta si paroma inverzna, ravno tako tudi ostanka 7 in 13. Ostanek 17 je znova inverz samemu sebi.

7.6.2 Linearne enačbe

Enačba $ax \equiv b \pmod{m}$ je *linearna kongruenčna enačba* z eno neznako v Z_m. Kaj lahko povemo o njeni rešljivosti in rešitvah? S prevedbo na linearne diofantske enačbe bomo pokazali naslednjo karakterizacijo.

Trditev 7.21 *Naj bosta $a, b \in Z_m$ in $a \neq 0$. Enačba*

$$a \cdot x \equiv b \pmod{m} \tag{7.38}$$

je rešljiva natanko tedaj, ko $\gcd(a,m) \mid b$. V tem primeru ima enačba v Z_m natanko $\gcd(a,m)$ rešitev.

Dokaz. Enačbo prepišemo v njeno linearno diofantsko ustreznico

$$a \cdot x + m \cdot y = b,$$

ki je po izreku 7.6 rešljiva natanko tedaj, ko $\gcd(a,m) \mid b$.

Če se omejimo na x-koordinate rešitev, potem po izreku 7.7 velja formula

$$x_t = x_0 + t \cdot \frac{m}{\gcd(a,m)}.$$

To pomeni, da je med števili $0, \ldots, m-1$ natanko $\gcd(a, m)$ števil-ostankov iz Z_m, ki rešijo enačbo (7.38). □

Oglejmo si zaporedje linearnih enačb

$$\begin{aligned} x &\equiv 3 \pmod{18} && (7.39)\\ 3x &\equiv 9 \pmod{18} && (7.40)\\ 9x &\equiv 9 \pmod{18} \\ 0x &\equiv 0 \pmod{18} \end{aligned}$$

Pridelali smo jih tako, da smo prvo izmed njih po vrsti pomnožili s 3, 9 in 0. Po trditvi 7.21 so vse enačbe rešljive in imajo po vrsti $\gcd(1, 18) = 1$, $\gcd(3, 18) = 3$, $\gcd(9, 18) = 9$ oziroma $\gcd(0, 18) = 18$ rešitev v Z_{18}. Vsaka rešitev posamezne od zgornjih enačb je tudi rešitev naslednje enačbe. Obratna implikacija pa, že zaradi števila rešitev, ne velja.

Fenomen poznamo v množici realnih števil. Če obe strani enačbe $x = 1$ pomnožimo z 0, potem število rešitev dramatično naraste, vsako realno število postane rešitev. Po drugi strani množenje enačbe $x = 1$ z neničelnim realnim številom ohrani družino rešitev.

V modulski aritmetiki opazimo vmesno verzijo takšnega fenomena. Ostanka 3 in 9 sta delitelja niča v Z_{18} in množenje (linearne) enačbe z deliteljem niča lahko poveča število rešitev enačbe. Prirastek števila rešitev pa je lahko odvisen od posameznega delitelja niča. Množenje linearne kongruenčne enačbe z 0 ima za posledico, da so vsi ostanki rešitve takšne enačbe.

Če enačbo (7.39) pomnožimo z *obrnljivim* elementom, denimo 5, potem množico rešitev ohranimo.

$$5x \equiv 15 \pmod{18} \qquad (7.41)$$

Enačba (7.41) ima natančno $\gcd(5, 18) = 1$ rešitev v Z_{18}. To je ostanek 3, ki je obenem tudi edina rešitev enačbe (7.39). Enačbo (7.39) lahko pridelamo iz enačbe (7.41), če jo pomnožimo z 11, ki je inverz za množenje od ostanka 5.

Razmišljanje lahko strnemo takole:

> *Množenje linearne kongruenčne enačbe z deliteljem niča lahko poveča množico rešitev. Množenje linearne kongruenčne enačbe z obrnljivim ostankom a množico rešitev ohrani, saj lahko originalno enačbo pomnožimo z inverzom a^{-1} in povrnemo prvotno enačbo.*

Kako se lotimo enačbe (7.40) v praksi? Prepišemo jo lahko v linearno diofantsko ustreznico

$$3 \cdot x + 18 \cdot y = 9. \qquad (7.42)$$

To enačbo lahko krajšamo s 3 in pridelamo

$$x + 6 \cdot y = 3, \tag{7.43}$$

ki v obratni smeri ustreza kongruenčni enačbi

$$x \equiv 3 \pmod 6 \tag{7.44}$$

Eno rešitev linearne diofantske enačbe (7.43), $x_0 = 3, y_0 = 0$, uganemo, vse rešitve pa zapišemo z enoparametrično družino kot

$$x_t = 3 + t \cdot 6$$
$$y_t = 0 - t \cdot 1.$$

To pomeni, da je $x = 3$ edina rešitev enačbe (7.44) v Z_6, oziroma da so $x_1 = 3, x_2 = 9$ in $x_3 = 15$ natančno vse tri rešitve enačbe (7.40) v Z_{18}.

7.6.3 Sistemi linearnih enačb

Omejili se bomo na sisteme linearnih kongruenčnih enačb z dvema neznankama, kjer sta obe enačbi kongruenčni enačbi po *istem* modulu m.

Sistem

$$\begin{array}{rcrcll} x & + & 2y & \equiv & 3 & \pmod{18} \\ 2x & + & 5y & \equiv & 13 & \pmod{18} \end{array} \tag{7.45}$$

bomo napadli z metodo *izrazi-in-vstavi*. Iz ene od enačb bomo izrazili eno od spremenljivk in jo vstavili v drugo.

V prvi enačbi se ponuja spremenljivka x. Če na obeh straneh kongruence odštejemo $2y$, pridelamo

$$x \equiv 3 - 2y \equiv 3 + 16y \pmod{18}. \tag{7.46}$$

Zvezo (7.46) vstavimo v drugo enačbo sistema (7.45) in pridelamo

$$2x + 5y \equiv 2(3 + 16y) + 5y \equiv 6 + 37y \equiv 6 + y \equiv 13 \pmod{18},$$

kar je enakovredno enačbi

$$y \equiv 13 - 6 \equiv 7 \pmod{18}. \tag{7.47}$$

Zvezo (7.47) uporabimo v (7.46) in računamo:

$$x \equiv 3 + 16y \equiv 3 + 16 \cdot 7 \equiv 7 \pmod{18}.$$

Torej je par $x = 7, y = 7$ *edina* rešitev sistema kongruenc (7.45) v Z_{18}.

Oglejmo si še en primer. Sistem

$$\begin{array}{rcrcll} 8x & + & 9y & \equiv & 10 & \pmod{18} \\ 3x & + & 6y & \equiv & 9 & \pmod{18} \end{array} \tag{7.48}$$

ima naslednjo težavo. Nobeden od koeficientov, s katerimi sta pomnoženi spremenljivki x in y, ni obrnljiv. V primeru obrnljivega koeficienta a v členu ax bi lahko ustrezno enačbo pomnožili z multiplikativnim inverzom a^{-1} in s tem osamili spremenljivko x.

Lahko pa sistem (7.48) enakovredno prepišemo v

$$\begin{array}{rcrcll} 8x & + & 9y & \equiv & 10 & \pmod{18} \\ 5x & + & 3y & \equiv & 1 & \pmod{18} \end{array} \tag{7.49}$$

ki smo ga dobili tako, da smo drugo enačbo prvotnega sistema nadomestili z razliko začetnih enačb.

Koeficient 5 je obrnljiv ostanek v Z_{18}, zato smemo drugo enačbo sistema pomnožiti z ostankom 11, ki je inverz ostanka 5 za množenje po modulu 18. Pridelamo

$$11 \cdot (5x + 3y) \equiv 55x + 33y \equiv x + 15y \equiv 11 \pmod{18}.$$

Odtod sledi

$$x \equiv 11 + 3y \pmod{18} \tag{7.50}$$

Vstavimo zvezo (7.50) v prvo enačbo sistema (7.49) in računajmo

$$8x + 9y \equiv 8 \cdot (11 + 3y) + 9y \equiv 16 + 15y \equiv 10 \pmod{18}$$

Zato je

$$15y \equiv 12 \pmod{18} \tag{7.51}$$

Enačba (7.51) ima po trditvi 7.21 natančno $\gcd(15, 18) = 3$ rešitve v Z_{18}. Z metodo (a) *prepiši kot linearno diofantsko enačbo*, (b) *krajšaj s* 3 in (c) *prepiši nazaj v kongruenčno enačbo*, pridelamo enakovredno kongruenčno enačbo

$$5y \equiv 4 \pmod{6}, \tag{7.52}$$

ki je v Z_6 enolično rešljiva. Rešitev $y \equiv 2 \pmod 6$ uganemo (ali pa izračunamo tako, da enačbo pomnožimo s 5), kar pridela tri rešitve enačbe (7.51) v Z_{18} in sicer

$$y_1 = 2, \qquad y_2 = 8, \qquad y_3 = 14. \tag{7.53}$$

Z vstavljanjem v zvezo (7.50) pa dobimo vse tri rešitve sistema (7.48)

$$\begin{aligned} x_1 &= 17, \quad y_1 = 2, \\ x_2 &= 17, \quad y_2 = 8, \\ x_3 &= 17, \quad y_3 = 14, \end{aligned}$$

v družini ostankov Z_{18}.

Poglavje 8

Permutacije

Preslikavi $f, g : A \to A$, ki slikata iz *iste* množice A nazaj v isto množico, znamo komponirati na dva načina. Tako $f \circ g$ kot $g \circ f$ sta preslikavi $A \to A$. Če sta f in g injektivni, potem sta tudi kompozituma injektivna, če sta preslikavi f in g surjektivni, sta takšna tudi kompozituma.

Če sta f in g celo bijekciji, potem sta tudi $f \circ g$ in $g \circ f$ bijekciji, obstajata pa tudi inverzni (bijektivni) preslikavi f^{-1} in g^{-1}.

V tem poglavju se bomo ukvarjali s *permutacijami*. Permutacija v množici A je, v širšem pomenu besede, bijektivna preslikava iz množice A vase.

Z računalniškega stališča imamo težavo pri neskončnih množicah A. Če se omejimo na eno bolj pohlevnih neskončnih množic, množico celih števil $\mathbb{Z}$, nam med bijekcijami na pamet pridejo le preslikave oblike $x \mapsto -x + 12$. Družina vseh bijekcij $\mathbb{Z} \to \mathbb{Z}$ pa je, verjemite mi na besedo, nepredstavljivo velika. Celo tako hudo je, da za tipično bijekcijo ni nobenega boljšega načina predstavitve kot z (neskončno) tabelico. Takšno (pa gre v resnici za eno sámo preslikavo) bi težko zapisali v računalniški pomnilnik.

Zato bomo za množico A, opazujemo bijekcije iz $A \to A$, privzeli, da je končna množica.

Se morda smemo tudi pri končnih množicah omejiti na kakšne prav posebne? Če opazujemo bijekcijo

$$f : \{\text{kumkvat}, \text{limona}, \text{pomaranča}, \text{grenivka}\} \to \{\text{kumkvat}, \text{limona}, \text{pomaranča}, \text{grenivka}\},$$

jo smemo opisati tako, da sadeže uredimo po velikosti (glej, glej, saj so že urejeni), potem pa za vsakega od citrusov ξ označimo, kateri sadež (po velikosti) je njegova funkcijska vrednost $f(\xi)$. Po domače, bijekcijo f nadomestimo z alternativno bijekcijo

$$f' : \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}.$$

Torej: Obravnavo bijekcij $A \to A$, kjer je A poljubna množica, bomo zožili na primer, ko je A končna množica. Ta omejitev je bistvene narave. Med končnimi množicami pa

se bomo omejili na množice naravnih števil $\{1, 2, \ldots, n\}$. To omejitev pa naredimo brez škode za splošnost.

8.1 Zapis permutacije

Permutacija dolžine n je bijektivna preslikava

$$\varphi : \{1, \ldots, n\} \to \{1, \ldots, n\}.$$

Dogovorimo se, da za imena permutacij uporabljamo male grške črke.

Družina vseh permutacij dolžine n je *simetrična grupa reda n* in jo označimo s S_n.

Permutacije bomo lahko zapisovali *s tabelico*. Za φ, permutacijo iz S_n, bomo v zgornji vrstici tabelice navedli števila od 1 do n, v spodnji vrstici tabelice pa njihove slike. Ker je permutacija bijektivna preslikava, se bo vsako od števil med 1 in n v spodnji vrstici pojavilo natančno enkrat.[1]

Zapišimo nekaj zgledov permutacij.

$$\begin{aligned}
\varphi_1 &= \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{smallmatrix}\right), \\
\varphi_2 &= \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 1 & 4 \end{smallmatrix}\right), \\
\varphi_3 &= \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 2 & 1 & 4 & 6 \end{smallmatrix}\right), \\
\varphi_4 &= \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 2 & 1 & 4 & 6 & 7 \end{smallmatrix}\right).
\end{aligned}$$

Permutacije $\varphi_1, \varphi_2, \varphi_3$ in φ_4 so po vrsti dolžin $4, 5, 6$ in 7. Permutacija φ_1 število 1 slika v 2, kar v funkcijskem zapisu predstavimo z $\varphi_1(1) = 2$. Velja tudi $\varphi_1(2) = 4$ in $\varphi_1(4) = 1$. Številu 3, za katerega velja $\varphi_1(3) = 3$, pravimo tudi *negibna točka* permutacije φ_1.

Opazimo pa še eno posebnost. Permutacijo $\varphi_3 \in S_6$ smo iz permutacije φ_2 dobili tako, da smo ji dodali stolpec s šestico tako zgoraj kot spodaj. Na podoben način smo pridelali tudi permutacijo φ_4.

Če permutacijo razumemo kot preslikavo, potem lako za φ_3 rečemo, da slika 6 v 6, $\varphi_3(6) = 6$. Z drugimi besedami, permutacija φ_3 šestico *pusti pri miru*. Toda tudi permutacija φ_2 število 6 *pusti pri miru*, saj 6 sploh ne spada v njeno definicijsko območje. Na preostalih manjših številih pa se φ_2 in φ_3 ujemata.

Tako lahko permutacijo *majhne dolžine* na naraven način interpretiramo kot permutacijo *večje dolžine*, ki bi jo dobili z dodajanjem stolpcev (z enakima številoma v obeh vrsticah).

Na naraven način lahko simetrično grupo reda n interpretiramo kot podmnožico simetrične grupe reda m, če je $m \geq n$. Velja torej veriga vsebovanosti

$$S_1 \subseteq S_2 \subseteq S_3 \subseteq S_4 \subseteq S_5 \subseteq \ldots \tag{8.1}$$

[1] Permutacije, v smislu preštevalne kombinatorike, so načini razporeditve družine različnih objektov v linearni vrstni red. Ustrezajo spodnjim vrsticam zapisov naših permutacij s tabelicami.

Permutacija *identitete* je permutacija *id*, ki vsako naravno število k slika v k. V skladu z (8.1) bomo identiteto dojemali kot pripadnico vseh simetričnih grup S_n, po domače, identiteta id je ena sama.

Če se bomo v obravnavi posebej omejili na, denimo, permutacije dolžine 4, potem bomo pisali

$$\mathrm{id} = \left(\begin{smallmatrix}1&2&3&4\\1&2&3&4\end{smallmatrix}\right) \in S_4.$$

Produkt permutacij in inverzna permutacija

Zgodbo o permutacijah bomo precej lažje vozili na konkretnih zgledih. Zapišimo torej permutaciji $\pi, \psi \in S_9$, pojavljali se bosta praktično do konca poglavja,

$$\begin{aligned}\pi &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\3&7&9&6&8&5&1&4&2\end{smallmatrix}\right),\\ \psi &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\2&4&6&1&3&5&7&9&8\end{smallmatrix}\right).\end{aligned} \tag{8.2}$$

Za množenje permutacij bomo uporabljali *relacijski produkt*. Dejstvi $\pi(1) = 3$ in $\psi(3) = 6$ lahko v relacijskem zapisu predstavimo kot $1\pi 3$ in $3\psi 6$. Odtod sledi $1(\pi * \psi)6$, kar interpretiramo kot $(\pi * \psi)(1) = 6$. Izračunajmo produkt $\pi * \psi$ do konca.

$$\begin{aligned}\pi * \psi &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\3&7&9&6&8&5&1&4&2\end{smallmatrix}\right) * \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\2&4&6&1&3&5&7&9&8\end{smallmatrix}\right)\\ &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\6&7&8&5&9&3&2&1&4\end{smallmatrix}\right)\end{aligned}$$

Izračunamo lahko tudi produkt $\psi * \pi$.

$$\begin{aligned}\psi * \pi &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\2&4&6&1&3&5&7&9&8\end{smallmatrix}\right) * \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\3&7&9&6&8&5&1&4&2\end{smallmatrix}\right)\\ &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\7&6&5&3&9&8&1&2&4\end{smallmatrix}\right)\end{aligned}$$

Produkta $\pi * \psi$ in $\psi * \pi$ nista enaka. Množenje permutacij *ni komutativno*, kot tudi množenje relacij in komponiranje preslikav nista komutativni operaciji.

Inverzno permutacijo zapišemo tako, da zamenjamo vrstici, nato pa stolpce uredimo po prvih koordinatah. Tako je

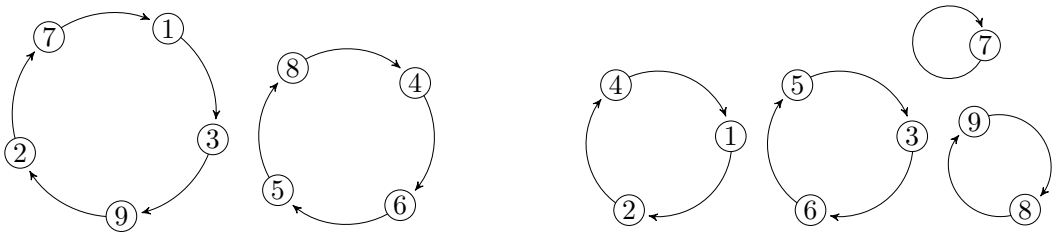
$$\begin{aligned}\pi^{-1} &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\7&9&1&8&6&4&2&5&3\end{smallmatrix}\right),\\ \psi^{-1} &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\4&1&5&2&6&3&7&9&8\end{smallmatrix}\right).\end{aligned} \tag{8.3}$$

Izračunajmo še produkt permutacije π z njej inverzno permutacijo π^{-1},

$$\begin{aligned}\pi * \pi^{-1} &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\3&7&9&6&8&5&1&4&2\end{smallmatrix}\right) * \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\7&9&1&8&6&4&2&5&3\end{smallmatrix}\right)\\ &= \left(\begin{smallmatrix}1&2&3&4&5&6&7&8&9\\1&2&3&4&5&6&7&8&9\end{smallmatrix}\right) \;=\; \mathrm{id}.\end{aligned}$$

Veljajo tudi enakosti

$$\pi^{-1} * \pi = \psi * \psi^{-1} = \psi^{-1} * \psi = \mathrm{id}.$$



Slika 8.1: Grafa permutacij $\pi = (1\,3\,9\,2\,7)(4\,6\,5\,8)$ in $\psi = (1\,2\,4)(3\,6\,5)(7)(8\,9)$.

Simetrična grupa S_n

V tem kratkem razdelku navedemo lastnosti, ki veljajo za družino vseh permutacij iz S_n.

Izrek 8.1 *Naj bodo $\varphi, \sigma, \rho \in S_n$. Potem*

(i) $\varphi * \sigma \in S_n$,

(ii) $\varphi^{-1} \in S_n$,

(iii) $\varphi * (\sigma * \rho) = (\varphi * \sigma) * \rho$,

(iv) $\varphi * \mathrm{id} = \mathrm{id} * \varphi = \varphi$,

(v) $\varphi * \varphi^{-1} = \varphi^{-1} * \varphi = \mathrm{id}$,

(vi) $\varphi^{-1} * \sigma^{-1} = (\sigma * \varphi)^{-1}$

Namesto dokaza komentar. Produkt permutacij je permutacija, ravno tako je inverz permutacije iz S_n znova permutacija iz S_n. Množenje permutacij je asociativno, saj je celo množenje relacij (bolj splošnih struktur) asociativno. Množenje z identiteto ne spremeni prvotnega izraza, tudi to lastnost poznamo že od relacij.

Produkt permutacije in njej inverzne permutacije je identiteta, formulo za produkt inverzov pa smo ravno tako že obdelali pri relacijah — spomni se trditve 4.1(ii).

Zapis z disjunktnimi cikli

Permutacije so posebne vrste preslikav in slednje so posebne vrste relacij. Na sliki 8.1 sta predstavljena grafa permutacij π in ψ (8.2).

Kaj lahko povemo o grafih permutacij? Za vsako število $k \in \{1, \ldots, n\}$ bo v grafu permutacije (a) natančno ena puščica *izstopala* iz k, ker je permutacija preslikava, in

(b) natančno ena puščica *vstopala* v k, ker je permutacija bijekcija. Sledenje zaporednim puščicam nas bo vedno in na enoličen način pripeljalo do začetne številke. Graf permutacije ima nekaj *disjunktnih* ciklov, ki so (če smo pedantni) usmerjeni.

Ciklom pravimo disjunktni, ker nobena številka ne nastopa hkrati v dveh ali več ciklih permutacije.

Graf permutacije π ima dva disjunktna cikla, ki sta dolžine 4 in 5. Graf permutacije ψ pa je sestavljen iz štirih disjunktnih ciklov, ki so po vrsti dolžine 3, 3, 2 in 1.

Ciklična struktura permutacije je informacija o številu in dolžinah ciklov, ki nastopajo v grafu permutacije (oziroma v zapisu z disjuktnimi cikli, ko ga spravimo pod streho).

Tako pravimo, da ima

> permutacija π ciklično strukturo $[4, 5]$ in
> permutacija ψ ciklično strukturo $[3, 3, 2, 1]$.

Vrstni red členov v zapisu ciklične strukture ni pomemben, tudi $[5, 4]$ je ciklična struktura permutacije π, saj je permutacija π *sestavljena* iz cikla dolžine 5 in cikla dolžine 4.

Ciklična struktura permutacije bo pomembno vplivala na algebrajske lastnosti permutacije — kako se permutacija obnaša v enačbah. Zato želimo uporabljati tudi alternativen način zapisa permutacije, iz katerega bo ciklična struktura takoj razvidna. To je *zapis z disjunktnimi cikli*. Permutaciji π in ψ v zapisu z disjunktnimi cikli zapišemo takole:

$$\begin{aligned} \pi &= (1\,3\,9\,2\,7)(4\,6\,5\,8), \\ \psi &= (1\,2\,4)(3\,6\,5)(7)(8\,9). \end{aligned} \tag{8.4}$$

Zapis $(1\,3\,9\,2\,7)$ cikla permutacije π preberemo na naslednji način. S permutacijo π se 1 slika v 3, 3 v 9, število 9 naprej v 2, 2 v 7 in končno 7 *nazaj* v 1.

Vrstni red ciklov v zapisu permutacije z disjunktnimi cikli ni pomemben. Prav tako ni pomembno, s katerim številom začnemo posamezen cikel zapisovati. Poleg tega bomo cikle dolžine 1 v zapisu permutacije z disjunktnimi cikli lahko izpuščali. Torej smemo permutaciji π in ψ zapisati tudi kot

$$\begin{aligned} \pi &= (1\,3\,9\,2\,7)(4\,6\,5\,8) = (4\,6\,5\,8)(1\,3\,9\,2\,7) \\ &= (3\,9\,2\,7\,1)(5\,8\,4\,6) = (7\,1\,3\,9\,2)(6\,5\,8\,4) \end{aligned}$$

in

$$\begin{aligned} \psi &= (1\,2\,4)(3\,6\,5)(7)(8\,9) = (8\,9)(4\,1\,2)(6\,5\,3)(7) \\ &= (1\,2\,4)(3\,6\,5)(8\,9) = (9\,8)(2\,4\,1)(5\,3\,6), \end{aligned}$$

pa še na kakšen drug način.

Vseeno bomo *priporočali* zapisa (8.4), ki ju konstruiramo takole. Začnemo s številom 1 in sledimo ciklu, ki vsebuje število 1. Vsak naslednji cikel pa začnemo zapisovati z

najmanjšim številom, ki ga v nastajajočem zapisu še nismo srečali. S takšnim pristopom denimo zagotovimo, da nobenega od ciklov permutacije pri zapisovanju ne izpustimo.

Še nekaj tèrminov za popotnico. Ciklu dolžine k bomo rekli tudi *k-cikel*. Posebej, 2-ciklu pravimo tudi *transpozicija*, 1-cikel pa je *negibna točka* permutacije.

Ciklična permutacija ali kar *cikel* je vsaka permutacija, ki ima v ciklični strukturi, razen morebitnih negibnih točk (ciklov dolžine 1), en sam cikel.

Naučimo se množiti permutacije, zapisane z disjunktnimi cikli. Pod drobnogled vzemimo produkt $\pi * \psi$.

$$\pi * \psi = (1\,3\,9\,2\,7)(4\,6\,5\,8) * (1\,2\,4)(3\,6\,5)(7)(8\,9) \ = (1\,6\,3\,8)(2\,7)(4\,5\,9)$$

Število a prenašamo preko ciklov proti desni, posamezen cikel nam število pusti na miru (če na primer a ne pripada trenutnemu ciklu) ali pa ga spremeni v njegovega naslednika a' (ki je lahko tudi prvo število zapisa posameznega cikla, če je a na koncu). Prenos preko ciklov nadaljujemo z a'. Ko s številom (morda se je vmes nekajkrat preobrazilo) prilezemo preko vseh ciklov obeh permutacij, ga zapišemo v zapis produkta.

Cikle produkta $\pi * \psi$ začnemo v skladu s priporočilom zapisovati z enico 1. Prvi cikel produkta vsebuje števila $1, 3, 6$ in 8, zato naslednji cikel v skladu s priporočilom začnemo zapisovati z 2.

Postopek množenja permutacij, zapisanih z disjunktnimi cikli, ne zahteva, da so cikli disjunktni. Z istim postopkom — prenosom števila preko ciklov — lahko izračunamo tudi produkt večjega števila permutacij. Pa še ena opazka, za posamezno permutacijo, zapisano z disjunktnimi cikli, lahko rečemo, da je enaka *produktu* svojih disjunktnih ciklov. Permutacija π je tako produkt (disjunktnih) 5-cikla in 4-cikla, permutacija ψ pa produkt (disjunktnih) 3-cikla, še enega 3-cikla in ene transpozicije.

Na tem mestu je jasno, zakaj smemo 1-cikle iz zapisa izpuščati. Rezultata množenja ne spremenijo, saj gre vsako od števil preko 1-cikla brez preobrazbe.

Zmnožimo permutaciji ψ in π še z drugim vrstnim redom faktorjev.

$$\psi * \pi = (1\,2\,4)(3\,6\,5)(7)(8\,9) \ * (1\,3\,9\,2\,7)(4\,6\,5\,8) = (1\,7)(2\,6\,8)(3\,5\,9\,4)$$

Permutaciji $\pi * \psi$ in $\psi * \pi$ imata isto ciklično strukturo $[2+3+4]$. To ni slučajno, temveč sistematično. Četudi produkt permutacij ni komutativen, je ciklična struktura produkta dveh permutacij neodvisna od vrstnega reda faktorjev.[2]

Kaj pa zapis inverzne permutacije z disjunktnimi cikli? Na najenostavnejši način gre takole. Zapis prepišemo od zadaj naprej. Če želimo, ga nato prepišemo ali v skladu s priporočilom ali na kakšen drug način in morda celo izpustimo 1-cikle. Tako je

$$\pi^{-1} = (8\,5\,6\,4)(7\,2\,9\,3\,1) = (1\,7\,2\,9\,3)(4\,8\,5\,6),$$
$$\psi^{-1} = (9\,8)(7)(5\,6\,3)(4\,2\,1) = (1\,4\,2)(3\,5\,6)(7)(8\,9).$$

[2] To boste morali verjeti na besedo.

Potenciranje permutacij

Za potenciranje permutacij bo zapis permutacij z disjunktnimi cikli bistveno primernejši kot zapis s tabelicami. Problem potenciranja permutacije, ki je zapisana z disjunktnimi cikli, bomo ločili na dva podproblema — (a) kakšna je interakcija med posameznimi cikli pri potenciranju in (b) kaj se dogaja s posameznim ciklom.

Trditev 8.2 *Naj bosta $\alpha_1, \alpha_2 \in S_n$ ciklični permutaciji, ki sta kot cikla disjunktna. Potem α_1 in α_2 komutirata,*

$$\alpha_1 * \alpha_2 = \alpha_2 * \alpha_1.$$

Dokaz. Z A_1 označimo množico števil cikla α_1 in z A_2 množico števil cikla α_2. Množico preostalih števil označimo z A_3, $A_3 = \{1, \ldots, n\} \setminus (A_1 \cup A_2)$. Dovolj je premisliti, da za vsako naravno število a, $1 \leq a \leq n$, velja enakost

$$(\alpha_1 * \alpha_2)(a) = (\alpha_2 * \alpha_1)(a).$$

Ločimo tri primere. Če $a \in A_1$, potem a ne pripada A_2 in zato velja $\alpha_2(a) = a$. Odtod pa sledi $(\alpha_2 * \alpha_1)(a) = \alpha_1(\alpha_2(a)) = \alpha_1(a)$. Po drugi strani pa tudi $\alpha_1(a)$ pripada množici A_1 in ne pripada A_2, zato je $(\alpha_1 * \alpha_2)(a) = \alpha_2(\alpha_1(a)) = \alpha_1(a)$.

S popolnoma analognim postopkom pokažemo, da tudi za $a \in A_2$ velja zveza $(\alpha_1 * \alpha_2)(a) = (\alpha_2 * \alpha_1)(a)$.

Če pa $a \in A_3$, potem je $\alpha_1(a) = \alpha_2(a) = a$ in zato tudi $(\alpha_1 * \alpha_2)(a) = a = (\alpha_2 * \alpha_1)(a)$. □

Z uporabo zadnje trditve pokažemo prvo polovico rezultata o potenciranju permutacij.

Trditev 8.3 *Naj bo $\varphi \in S_n$ in*

$$\varphi = \alpha_1 * \alpha_2 * \cdots * \alpha_\ell$$

zapis permutacije φ z disjunktnimi cikli — permutacije $\alpha_1, \ldots, \alpha_\ell$ so natančno vsi disjunktni cikli permutacije φ. Potem za vsak eksponent $k \in \mathbb{N}$ velja

$$\varphi^k = \alpha_1^k * \alpha_2^k * \cdots * \alpha_\ell^k$$

Dokaz. Zapišimo

$$\varphi^k = (\alpha_1 * \alpha_2 * \cdots * \alpha_\ell)^k$$

in potenco produkta razpišimo. Po trditvi 8.2 vsak par ciklov α_i in α_j komutira (tudi, če je $i = j$). Zato lahko faktorje v razpisanem produktu uredimo po naraščajočih indeksih. □

Potrebno je samo še izračunati, kaj se dogaja s potencami posameznega cikla. V ta namen bomo potencirali cikla dolžine 5 oziroma 6

$$\alpha = (1\,2\,3\,4\,5) \quad \text{in} \quad \beta = (1\,2\,3\,4\,5\,6).$$

Standardna definicija potence z eksponentom 0 je $\alpha^0 = \beta^0 = \mathrm{id}$.

Začnimo s permutacijo α.

$$\begin{aligned}
\alpha^1 &= \alpha = (1\,2\,3\,4\,5)\\
\alpha^2 &= \alpha * \alpha = (1\,2\,3\,4\,5) * (1\,2\,3\,4\,5) = (1\,3\,5\,2\,4)\\
\alpha^3 &= \alpha^2 * \alpha = (1\,3\,5\,2\,4) * (1\,2\,3\,4\,5) = (1\,4\,2\,5\,3)\\
\alpha^4 &= \alpha^3 * \alpha = (1\,4\,2\,5\,3) * (1\,2\,3\,4\,5) = (1\,5\,4\,3\,2) = (5\,4\,3\,2\,1)\\
\alpha^5 &= \alpha^4 * \alpha = (5\,4\,3\,2\,1) * (1\,2\,3\,4\,5) = (1)(2)(3)(4)(5) = \mathrm{id}
\end{aligned}$$

Če je eksponent $k > 5$, potem lahko pišemo

$$\alpha^k = \alpha^5 * \alpha^{k-5} = \mathrm{id} * \alpha^{k-5} = \alpha^{k-5},$$

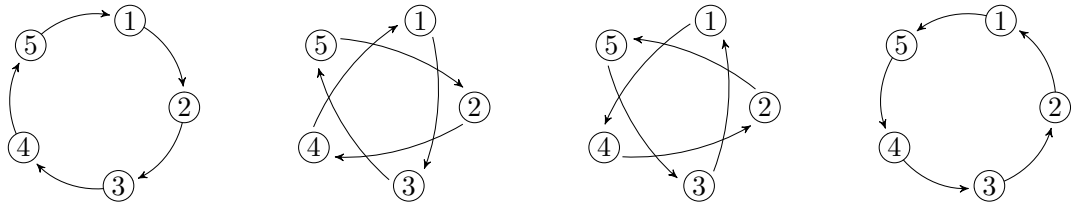
in zato induktivno sledi

$$\alpha^k = \alpha^{k \bmod 5}.$$

Še inverz — iz $\alpha^4 * \alpha = \alpha * \alpha^4 = \mathrm{id}$ $(= \alpha^5)$ sledi, da je

$$\alpha^{-1} = \alpha^4.$$

Na sliki 8.2 so prikazani grafi permutacij $\alpha, \alpha^2, \alpha^3$ in α^4.



Slika 8.2: 5-cikel $\alpha = (1\,2\,3\,4\,5)$ in njegove potence α^2, α^3 in α^4.

Kaj pa potence permutacije β? Računajmo.

$$\begin{aligned}
\beta^1 &= \beta = (1\,2\,3\,4\,5\,6)\\
\beta^2 &= \beta * \beta = (1\,2\,3\,4\,5\,6) * (1\,2\,3\,4\,5\,6) = (1\,3\,5)(2\,4\,6)\\
\beta^3 &= \beta^2 * \beta = (1\,3\,5)(2\,4\,6) * (1\,2\,3\,4\,5\,6) = (1\,4)(2\,5)(3\,6)\\
\beta^4 &= \beta^3 * \beta = (1\,4)(2\,5)(3\,6) * (1\,2\,3\,4\,5\,6) = (1\,5\,3)(2\,6\,4)\\
\beta^5 &= \beta^4 * \beta = (1\,5\,3)(2\,6\,4) * (1\,2\,3\,4\,5\,6) = (1\,6\,5\,4\,3\,2) = (6\,5\,4\,3\,2\,1)\\
\beta^6 &= \beta^5 * \beta = (6\,5\,4\,3\,2\,1) * (1\,2\,3\,4\,5\,6) = (1)(2)(3)(4)(5)(6) = \mathrm{id}
\end{aligned}$$

Reciklirajmo idejo. Če je eksponent $k > 6$, potem lahko pišemo

$$\beta^k = \beta^6 * \beta^{k-6} = \mathrm{id} * \beta^{k-6} = \beta^{k-6},$$

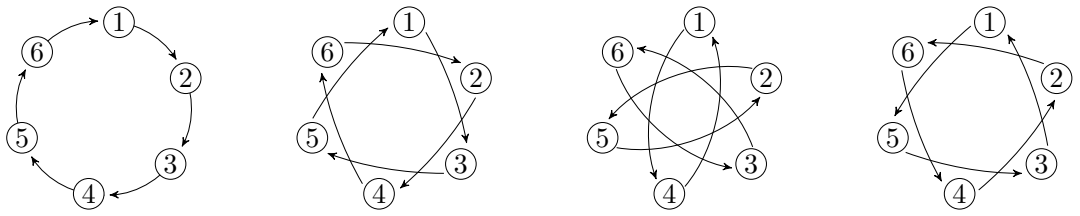
in zato induktivno sledi

$$\beta^k = \beta^{k \bmod 6}.$$

Inverz permutacije β pa je enak njeni peti potenci, iz $\beta^5 * \beta = \beta * \beta^5 = \mathrm{id}\ (= \beta^6)$ sledi, da je

$$\beta^{-1} = \beta^5.$$

Grafi nekaj začetnih potenc permutacije β so prikazani sliki 8.3.



Slika 8.3: 6-cikel $\beta = (1\,2\,3\,4\,5\,6)$ in njegove potence β^2, β^3 in β^4.

Kaj pa ciklična struktura potenc? Potence 5-cikla α imajo ali ciklično strukturo [5] (en sam cikel dolžine 5) ali pa so enake identiteti, in imajo ciklično strukturo $[1, 1, 1, 1, 1]$.

Ciklične strukture potenc permutacije β so bolj raznovrstne. Odvisno od eksponenta so to lahko [6] in $[1, 1, 1, 1, 1, 1]$, kot v primeru 5-cikla α, pridelamo pa tudi dodatni ciklični strukturi $[2, 2, 2]$ in $[3, 3]$.

Kar sami se zastavljata naslednji vprašanji: Kakšna je razlika med številoma 5 in 6, da prihaja do teh razlik? Je možno pri potenciranju cikla pridelati permutacijo, katere ciklična struktura vsebuje cikle različnih dolžin?

Izberimo število $a \in \{1, 2, 3, 4, 5\}$. Kaj so

$$\alpha(a), \alpha^2(a), \alpha^3(a), \ldots \quad \text{in} \quad \beta(a), \beta^2(a), \beta^3(a), \ldots?$$

Odgovor smo spoznali pri relacijskih potencah, trditvi 4.2. Za vsako naravno število k je $\alpha^k(a)$ število, ki ga dosežemo s sprehodom dolžine k v grafu permutacije α, če začnemo v a. Ker je graf permutacije α cikel, je sprehod, če poznamo začetek in dolžino, enolično določen.

Sprehod za 5 korakov vzdolž cikla α nas pripelje nazaj v začetno točko. Zato je $\alpha^5 = \mathrm{id}$. Podobno sprehod dolžine 4 ustreza sprehodu v napačno smer dolžine 1. Zato je $\alpha^{-1} = \alpha^4$. Končno lahko sprehod dolžine $5 \cdot k + r$ v grafu permutacije α reduciramo po modulu 5, saj sprehod dolžine $5 \cdot k$ pomeni vrnitev v začetek.

S temi opazkami smo pripravljeni razložiti dinamiko ciklične strukture pri potenciranju ciklične permutacije.

Trditev 8.4 *Naj bo α ciklična permutacija in n dolžina edinega njenega cikla. Potem ima permutacija α^k v zapisu z disjunktnimi cikli natančno* $\gcd(n, k)$ *ciklov, ki so vsi iste dolžine* $n/\gcd(n, k)$.

Dokaz. Zapišimo permutacijo α z n-ciklom,

$$\alpha = (a_0\, a_1\, a_2\, \ldots\, a_{n-2}\, a_{n-1}),$$

in izberimo njegovega pripadnika a_τ. V zapisu cikla γ_τ permutacije α^k (če $a_\tau = a_{(\tau+0\cdot k) \bmod n}$ zapišemo na prvem mestu) si števila sledijo v naslednjem zaporedju

$$a_{(\tau+0k) \bmod n},\ a_{(\tau+1k) \bmod n},\ a_{(\tau+2k) \bmod n},\ a_{(\tau+3k) \bmod n},\ a_{(\tau+4k) \bmod n},\ \ldots$$

Cikel γ_τ je dolžine x_0 natanko tedaj, ko je x_0 najmanjša pozitivna x-koordinata rešitve linearne diofantske enačbe

$$k \cdot x + n \cdot y = 0. \tag{8.5}$$

V tem primeru namreč velja $a_\tau = a_{(\tau+x_0\cdot k) \bmod n}$. Enačba (8.5) je zagotovo rešljiva, saj je njena desna stran enaka 0. Člen $k \cdot x_0$ pa je najmanjši možen natanko tedaj, ko je $k \cdot x_0 = \mathrm{lcm}(k, n)$. Po izreku 7.5 je v tem primeru $x_0 = n/\gcd(k, n)$.

Dolžina cikla γ_τ je neodvisna od izbire elementa a_τ. Zato so disjunktni cikli permutacije α^k vsi iste dolžine $n/\gcd(n, k)$. Njihovo število pa je enako $\gcd(n, k)$, saj kumulativno vsebujejo natanko vsa števila cikla α. □

Potenciranje posameznega cikla strnimo v še eno trditev.

Trditev 8.5 *Naj bo α ciklična permutacija in n dolžina njenega edinega cikla. Potem je dolžina cikla n najmanjši strogo pozitivni eksponent, za katerega je ustrezna potenca permutacije α enaka* id*,*

$$n = \min\{m \mid m > 0 \textit{ in } \alpha^m = \mathrm{id}\}.$$

Nadalje lahko inverz permutacije α izrazimo z

$$\alpha^{-1} = \alpha^{n-1},$$

velja pa tudi zveza

$$\alpha^k = \alpha^{k \bmod n}.$$

Dokaz. Iščemo najmanjši strogo pozitivni eksponent m, pri katerem so disjunktni cikli permutacije α^m vsi dolžine 1. Po trditvi 8.4 mora veljati $\gcd(m, n) = n$. Odtod sledi $m = n$.

Ker se permutacija α^{n-1} obnaša tako kot inverz α^{-1}, velja namreč

$$\alpha^{n-1} * \alpha = \alpha * \alpha^{n-1} = \mathrm{id},$$

po izreku 5.9 velja $\alpha^{-1} = \alpha^{n-1}$.

Preostanek trditve o redukciji eksponenta po modulu dolžine cikla enostavno sledi iz zveze $\alpha^n = \mathrm{id}$. □

Končno lahko zapišemo zaključni izrek o potenciranju permutacij, ki je enostavna posledica trditev 8.3 in 8.5.

Izrek 8.6 *Naj bo $\varphi \in S_n$ in*

$$\varphi = \alpha_1 * \cdots * \alpha_\ell$$

zapis permutacije φ z disjunktnimi cikli, ki so po vrsti dolžin $n_1, \ldots, n_\ell$. Potem je

$$\varphi^k = \alpha_1^{k \bmod n_1} * \cdots * \alpha_\ell^{k \bmod n_\ell}.$$

Kot zgled izračunajmo permutaciji π^{123} in ψ^{321}.

$$\begin{aligned}
\pi^{123} &= ((1\,3\,9\,2\,7)(4\,6\,5\,8))^{123} \\
&= (1\,3\,9\,2\,7)^{123}\,(4\,6\,5\,8)^{123} \\
&= (1\,3\,9\,2\,7)^{123 \bmod 5}\,(4\,6\,5\,8)^{123 \bmod 4} \\
&= (1\,3\,9\,2\,7)^{3}\,(4\,6\,5\,8)^{3} \\
&= (1\,2\,3\,7\,9)(4\,8\,5\,6)
\end{aligned}$$

$$\begin{aligned}
\psi^{321} &= ((1\,2\,4)(3\,6\,5)(7)(8\,9))^{321} \\
&= (1\,2\,4)^{321}\,(3\,6\,5)^{321}\,(7)^{321}\,(8\,9)^{321} \\
&= (1\,2\,4)^{321 \bmod 3}\,(3\,6\,5)^{321 \bmod 3}\,(7)^{321 \bmod 1}\,(8\,9)^{321 \bmod 2} \\
&= (1\,2\,4)^{0}\,(3\,6\,5)^{0}\,(7)^{0}\,(8\,9)^{1} \\
&= (8\,9)
\end{aligned}$$

Red permutacije φ, $\operatorname{ord}(\varphi)$, je najmanjše pozitivno naravno število k, za katerega je $\varphi^k = \mathrm{id}$. Če je φ ciklična permutacija, potem je $\operatorname{ord}(\varphi)$ enak njeni dolžini. Za identiteto id velja $\operatorname{ord}(\mathrm{id}) = 1$ in identiteta je *edina* permutacija reda 1.

Trditev 8.7 *Naj bo $\varphi \in S_n$. Njen red* $\operatorname{ord}(\varphi)$ *je najmanjši skupni večkratnik dolžin ciklov v zapisu permutacije φ z disjunktnimi cikli.*

Dokaz. Če velja $\varphi^k = \mathrm{id}$, potem je k večkratnik dolžine vsakega cikla α, ki nastopa v zapisu permutacije φ z disjunktnimi cikli. Najmanjši možen eksponent, pri katerem dobimo identiteto, je torej najmanjši skupni večkratnik dolžin ciklov iz zapisa. □

Izračunajmo reda permutacij π in ψ. Ciklična struktura permutacije π je $[4+5]$, zato je $\operatorname{ord}(\pi) = \operatorname{lcm}(4,5) = 20$. Ciklična struktura permutacije ψ je enaka $[3+3+2+1]$, njen red $\operatorname{ord}(\psi)$ je enak $\operatorname{lcm}(3,3,2,1) = 6$.

8.2 Parnost permutacij

Vsako naravno število $n \geq 2$ lahko zapišemo kot produkt praštevil in vsako naravno število $n \geq 1$ lahko zapišemo kot vsoto enic. Bistvo teh trditev se skriva v dejstvu, da lahko objekte velike množice zgradimo z uporabo algebrskih operacij in omejenega

nabora osnovnih gradnikov. Slednjih je *precej manj* (na tem mestu tèrmina *precej manj* ne bi definirali matematično) kot objektov v celi družini.

Tudi za permutacije velja podobna lastnost. Pokazali bomo, da lahko vsako permutacijo zapišemo kot produkt transpozicij (ne nujno disjunktnih). Enoličnosti pa ne bomo mogli zahtevati, posamezno permutacijo bomo znali zapisati kot produkt transpozicij na veliko različnih načinov.

Začnimo z permutacijo id. Transpozicija (cikel dolžine 2) τ je permutacija reda 2. To pomeni, da je $\tau^2 = \mathrm{id}$ oziroma da permutacijo identitete lahko zapišemo kot produkt dveh enakih (a poljubnih) transpozicij.

Zapišimo nekaj alternativ.

$$\begin{aligned}\mathrm{id} &= (1\,2)(1\,2) = (3\,4)(3\,4) = (1\,2)(1\,2)(1\,3)(1\,3)\\ &= (1\,2)(3\,4)(1\,2)(3\,4) = (1\,3)(1\,3)(1\,3)(1\,3)\end{aligned} \tag{8.6}$$

Izrek 8.8 *Vsako permutacijo $\varphi \in S_n$, $n \geq 2$, lahko zapišemo kot produkt transpozicij.*

Dokaz. Identiteto smo kot produkt transpozicij zapisali v (8.6). Vsaj prvi od omenjenih zapisov je ustrezen tudi v S_2.

Poljubna permutacija $\varphi \neq \mathrm{id}$ ima v svoji ciklični strukturi vsaj en cikel dolžine ≥ 2. Dovolj bo pokazati, da lahko vsak dolg cikel zapišemo kot produkt transpozicij.

Cikel dolžine 2 je transpozicija sam zase. Cikle večjih dolžin pa lahko zapišemo kot produkt transpozicij po naslednjem *receptu*.

$$(1\,2\,3\,4\,5\,6\,7) \;=\; (1\,2)(1\,3)(1\,4)(1\,5)(1\,6)(1\,7) \tag{8.7}$$

Enakost lahko preverimo z direktnim računom.

V zapisu permutacije z disjunktnimi cikli odstranimo negibne točke, cikle dolžine ≥ 3 pa prepišemo kot produkt transpozicij v skladu z (8.7). □

Kot zgled zapišimo permutaciji π in ψ kot produkt transpozicij.

$$\begin{aligned}\pi &= (1\,3\,9\,2\,7)(4\,6\,5\,8) = (1\,3)(1\,9)(1\,2)(1\,7)(4\,6)(4\,5)(4\,8)\\ \psi &= (1\,2\,4)(3\,6\,5)(7)(8\,9) = (1\,2)(1\,4)(3\,6)(3\,5)(8\,9)\end{aligned} \tag{8.8}$$

Zapis permutacije kot produkt transpozicij ni enoličen, to smo spoznali na primeru identitete (8.6). Tudi uporaba *recepta* (8.7) ne bo pridelala enoličnega zapisa, saj lahko posamezno permutacijo zapišemo kot produkt disjunktnih ciklov na različne načine — če spreminjamo vrstni red ciklov oziroma začetna števila v zapisih ciklov.

Vseeno zapis permutacije kot produkt transpozicij ni čisto brez strukture. Zdi se, da identitete id ne moremo zapisati kot produkt lihega števila transpozicij. Vsi naši produkti (8.6) uporabijo sodo mnogo faktorjev.

Fenomen je strukturne narave. Čeprav lahko posamezno permutacijo φ zapišemo kot produkt transpozicij na veliko različnih načinov, bodo vsi produkti uporabili bodisi sodo mnogo faktorjev bodisi bo v vsakem od produktov liho mnogo faktorjev. Velja naslednji izrek.

Izrek 8.9 (o parnosti permutacij) *Naj bo $\varphi \in S_n$, $n \geq 2$, in denimo, da π zapišemo kot produkt transpozicij na dva načina*

$$\pi = \tau_1 * \tau_2 * \cdots * \tau_k \quad in$$
$$\pi = \tau'_1 * \tau'_2 * \cdots * \tau'_\ell \,,$$

pri čemer so $\tau_1, \ldots, \tau_k, \tau'_1, \ldots, \tau'_\ell$ transpozicije. Potem je

$$k \equiv \ell \pmod 2.$$

Dokaz izreka o parnosti odložimo, trditev sámo pa vseeno uporabimo za naslednjo definicijo. Permutacija φ je *soda*, če jo lahko zapišemo kot produkt sodega števila transpozicij. Permutacija φ je *liha*, če jo lahko zapišemo kot produkt lihega števila transpozicij.

Izrek 8.8 trdi, da je vsaka permutacija soda ali liha, saj vsako lahko zapišemo kot produkt transpozicij, ki je sodo ali liho število. Izrek 8.9 pa zapiše, da permutacija ne more biti soda in liha hkrati. Torej je *parnost* vsake permutacije dobro in smiselno določena.

Parnost celih števil se lepo ujame s seštevanjem. Parnost vsote dveh števil je odvisna le od parnosti členov. Pri permutacijah se parnost lepo ujame s produktom (druge operacije tako ali tako nimamo na voljo). Produkt dveh permutacij iste parnosti bo soda permutacija, če pa zmnožimo dve permutaciji različnih parnosti, bo njun produkt liha permutacija. Utemeljitev je sila preprosta, vsakega od faktorjev zapišemo kot produkt transpozicij.

Parnost permutacije bomo ugnali okrog ovinka. Izmerili bomo, kako *oddaljena* je permutacija od permutacije identitete. Identiteta ima lastnost, da so v spodnji vrstici njenega zapisa s tabelico števila urejena v *pravem vrstnem redu*.

Naj bo $\varphi \in S_n$. Par števil i, j, $1 \leq i < j \leq n$, je *v inverziji* v permutaciji φ, če se v spodnji vrstici zapisa permutacije φ s tabelico števili i in j pojavita v *napačnem* vrstnem redu: večje število j je zapisano bolj levo od manjšega števila i.

V permutaciji

$$\pi = \left(\begin{smallmatrix} 1&2&3&4&5&6&7&8&9 \\ 3&7&9&6&8&5&1&4&2 \end{smallmatrix}\right)$$

so pari števil $1, 3$ in $2, 4$ in $5, 8$ v inverziji. Para $6, 8$ in $7, 9$ pa nista v inverziji.

Permutacija ψ je pohlevnejša kar se tiče inverzij. Tako malo jih je, da jih lahko celo naštejemo. V permutaciji

$$\psi = \left(\begin{smallmatrix} 1&2&3&4&5&6&7&8&9 \\ 2&4&6&1&3&5&7&9&8 \end{smallmatrix}\right)$$

je natančno sedem inverzij. To so naslednji pari: $1, 2$; $1, 4$; $1, 6$; $3, 4$; $3, 6$; $5, 6$ in $8, 9$.

Število inverzij v permutaciji φ označimo z $\mathrm{inv}(\varphi)$. Lahko si predstavljamo, da število inverzij permutacije meri, kako močno se dotična permutacija razlikuje od identitete. Za permutacijo ψ velja $\mathrm{inv}(\psi) = 7$. Pri tem je število inverzij s permutacijo sámo enolično določeno, medtem ko število transpozicij v produktu transpozicij, s katerim permutacijo izrazimo, ni enolično.

Trditev 8.10 *Denimo, da je φ permutacija in τ transpozicija. Potem se števili inverzij v permutacijah π in $\pi * \tau$* razlikujeta *po parnosti,*

$$\mathrm{inv}(\varphi) \not\equiv \mathrm{inv}(\varphi * \tau) \pmod 2.$$

Dokaz. Naj bo $\tau = (p\,q)$, zapišimo pa tudi permutaciji φ in $\varphi * \tau = \varphi * (p\,q)$ v tabelarični obliki. Na tem mestu lahko brez škode za splošnost privzamemo, da je v spodnji vrstici permutacije φ število p zapisano levo od števila q.

$$\varphi = \begin{pmatrix} 1 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & n \\ \boxed{\ell_1, \ell_2, \ell_3, \dots} & & & & p & & & & q & & & & \boxed{d_1, d_2, d_3, \dots} \end{pmatrix}$$

$$\varphi * (p\,q) = \begin{pmatrix} 1 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & n \\ \boxed{\ell_1, \ell_2, \ell_3, \dots} & & & & q & & & & p & & & & \boxed{d_1, d_2, d_3, \dots} \end{pmatrix}$$

Zapisa permutacij φ in $\varphi * (p\,q)$ se razlikujeta samo v mestih števil p in q. Vsa ostala števila v obeh tabelicah so zapisana v natančno istem vrstnem redu. Lokacije preostalih števil glede na mesti p in q v tabelici ločimo na *levo*, kjer ležijo števila $\ell_1, \ell_2, \ell_3, \dots$, *desno* s števili $d_1, d_2, d_3, \dots$, in *center*, kjer ležijo $c_1, c_2, c_3, \dots$. Opazovali bomo spremembo v inverzijah glede na permutaciji φ in $\varphi * (p\,q)$, natančneje spremembe relativnih položajev parov števil v obeh permutacijah.

(1) *K spremembi števila inverzij lahko prispevajo samo pari števil a, b, kjer je vsaj eden od a, b enak p ali q.*

Če sta tako a in b različna od p in q, potem je njuna relativna lega v φ in $\varphi * (p\,q)$ enaka. Torej je par a, b v inverziji v obeh ali pa v nobeni od obeh permutacij.

(2) *Leva števila $\ell_1, \ell_2, \ell_3, \dots$ ne prispevajo k spremembi števila inverzij.*

Po (1) opazujemo samo pare a, b, kjer je a eno od levih števil in b eden od p, q. Toda a je v obeh permutacijah φ in $\varphi * (p\,q)$ levo od b in njuna relativna položaja se ne spremenita.

Isti premislek lahko naredimo s števili desno.

(3) *Tudi desna števila $d_1, d_2, d_3, \dots$ ne prispevajo k spremembi števila inverzij.*

Centralna števila $c_1, c_2, c_3, \dots$ razdelimo na

majhna, $m_1, m_2, m_3, \ldots$, strogo manjša od $\min(p, q)$,
velika, $v_1, v_2, v_3, \ldots$, strogo večja od $\max(p, q)$ in
srednja, $s_1, s_2, s_3, \ldots$, strogo med $\min(p, q)$ in $\max(p, q)$.

(4) *Majhna števila $m_1, m_2, m_3, \ldots$ ne prispevajo k spremembi števila inverzij.*

Če je a majhno število, je v permutaciji π v inverziji z p in ni v inverziji s q. V permutaciji $\varphi * (p\,q)$ pa je ravno obratno. Par a, q je in par a, p ni v inverziji v permutaciji $\varphi * (p\,q)$.

Idejo recikliramo tudi na velikih številih.

(5) *Velika števila $v_1, v_2, v_3, \ldots$ ne prispevajo k spremembi števila inverzij.*

Pri sredinskih številih je malo drugače. Če je a sredinsko število, je a v natanko eni od permutacij $\varphi, \varphi * (p\,q)$ v inverziji z *obema* p in q, v drugi pa v inverziji z *nobenim* od p, q (odvisno od tega, kateri od p, q je večji). Sredinsko število sicer prispeva k spremembi števila inverzij, toda parnosti ne pokvarimo.

(6) *Sredinska števila $s_1, s_2, s_3, \ldots$ ne prispevajo k* spremembi parnosti števila *inverzij.*

Ostal nam je samo par p, q. Ti dve števili sta v inverziji v natančno eni od permutacij $\varphi, \varphi * (p\,q)$. Par p, q torej prispeva ± 1 k spremembi števila inverzij. Ob upoštevanju (1),. . . ,(6), je dokaz trditve 8.10 zaključen. □

Na tem mestu lahko dokončamo dokaz izreka o parnosti.

Dokaz.[izreka 8.9] Naj bo φ poljubna permutacija, zapišimo jo kot produkt transpozicij

$$\varphi = \mathrm{id} * \tau_1 * \tau_2 * \ldots * \tau_\ell,$$

s prikladnim faktorjem id na skrajni levi strani. Dovolj je premisliti, da velja $\mathrm{inv}(\varphi) \equiv \ell \pmod 2$.

Definirajmo zaporedje permutacij φ_i, $i = 0, \ldots, \ell$, z naslednjim rekurzivnim opisom

$$\begin{aligned} \varphi_0 &= \mathrm{id}, \\ \varphi_i &= \varphi_{i-1} * \tau_i, \quad \text{če je } i \geq 1. \end{aligned}$$

Pri tem za zadnji člen zaporedja velja $\varphi_\ell = \varphi$.

Po trditvi 8.10 za vsak $i = 1, \ldots, \ell$ velja $\mathrm{inv}(\varphi_i) \not\equiv \mathrm{inv}(\varphi_{i-1}) \pmod 2$. Ker je $\mathrm{inv}(\varphi_0) \equiv \mathrm{inv}(\mathrm{id}) \equiv 0 \pmod 2$, lahko induktivno pokažemo veljavnost kongruence

$$\mathrm{inv}(\varphi_i) \equiv i \pmod 2, \quad \text{za vsak } i = 0, \ldots, \ell.$$

To pomeni, da je

$$\mathrm{inv}(\varphi) \equiv \mathrm{inv}(\varphi_\ell) \equiv \ell \pmod 2$$

in dokaz je končan. □

Enostavna posledica zgornjega dokaza je naslednja trditev.

Trditev 8.11 *Permutacija $\varphi \in S_n$ je soda natanko tedaj, ko je* $\mathrm{inv}(\varphi)$ *sodo število, in je liha natanko tedaj, ko je* $\mathrm{inv}(\varphi)$ *liho število.*

Permutaciji π in ψ smo kot produkt transpozicij izrazili v (8.8). Permutaciji sta lihi, v produktih imamo namreč 7 oziroma 5 faktorjev.

Število inverzij permutacije ψ, $\mathrm{inv}(\psi) = 7$, se po parnosti ujema s parnostjo permutacije ψ, kar konec koncev pričakujemo po trditvi 8.11. Števila $\mathrm{inv}(\pi)$ nismo izračunali, vemo pa, po isti trditvi, da je ravno tako liho.

Vrnimo se k prvotni motivaciji razdelka: Izraziti vsako permutacijo z ustrezno manjšim številom gradnikov. Uspelo nam je s transpozicijami. Se da celo bolje? Brez dokaza navedimo, da lahko vsako permutacijo φ izrazimo kot produkt transpozicij oblike

$$(1\,2), (1\,3), (1\,4), \ldots, (1\,n),$$

pa tudi kot produkt transpozicij oblike

$$(1\,2), (2\,3), (3\,4), \ldots, (n-1\,n).$$

8.3 Potenčna enačba s permutacijami

Permutacijska potenčna enačba je enačba oblike

$$\varphi^k = \alpha, \tag{8.9}$$

kjer je α znana permutacija, φ neznana permutacija in $k \in \mathbb{N}$ izbran eksponent. Morebitni negativni eksponent lahko v pozitivnega spremenimo tako, da bodisi nadomestimo spremenljivko φ z novo, njenim inverzom, bodisi pa invertiramo obe strani enakosti.

Tipična vprašanja bodo naslednja:

> Je enačba rešljiva?
> Lahko poiščeš kakšno rešitev enačbe? Morda nekaj bistveno različnih?
> Znaš poiskati vse rešitve enačbe?

Pri tem bo zadnje vprašanje v splošnem predstavljalo pretrd oreh. Zadovoljni bomo z odgovori na preostala vprašanja.

Za majhne vrednosti eksponenta k je analiza enostavna. Če je $k = 0$, potem je enačba (8.9) rešljiva natanko tedaj, ko je desna stran $\alpha = \mathrm{id}$. Res pa je, da so v tem primeru *vse* permutacije iz S_n rešitve. Če je $k = 1$, potem je enačba pri vsakem α enolično rešljiva.

Zgodba postane zanimiva z eksponentom $k \geq 2$. Za začetek uporabimo idejo o parnosti permutacij. Permutacija π^2 je soda, ne glede na to, kakšna je permutacija π. Zato enačba $\varphi^2 = \alpha$ nima rešitev, če je α liha permutacija. Če je α soda permutacija, pa odgovor ne bo niti enoznačen niti enostaven.

Zapis permutacije z disjunktnimi cikli smo motivirali z algebrajskimi lastnostmi permutacije. Pri rešljivosti potenčne enačbe (8.9) bo ključno vlogo odigrala ciklična struktura permutacije α.

Naj bosta ξ in ζ ciklični strukturi permutacij iz S_n. Ciklična struktura permutacije π^k je odvisna od eksponenta k in ciklične strukture permutacije π. Če velja, da ima π ciklično strukturo ξ in π^k ciklično strukturo ζ, to označimo z

$$\xi \xrightarrow{k} \zeta.$$

Zapis

$$\xi \to \zeta$$

pa pomeni, da obstaja eksponent m, za katerega je $\xi \xrightarrow{m} \zeta$. Navedimo nekaj zgledov, utemeljitev sledi iz trditve 8.4.

$$\begin{array}{lll} [6] \xrightarrow{2} [3,3], & [6,2] \xrightarrow{3} [2,2,2,2], & [2,3,4] \xrightarrow{7} [2,3,4], \\ [4,2] \xrightarrow{2} [2,2,1,1], & [4,2] \xrightarrow{3} [4,2], & [6,8] \xrightarrow{4} [3,3,2,2,2,2]. \end{array}$$

Denimo, da je ζ_α ciklična struktura znane permutacije α v enačbi (8.9). Ciklična struktura ξ je *dopustna*, če velja $\xi \xrightarrow{k} \zeta_\alpha$. Naslednjo trditev bomo dokazali z uporabo *metode nedoločenih koeficientov*.

Trditev 8.12 *Naj bo $\varphi^k = \alpha$ potenčna enačba z znano permutacijo α, ki ima ciklično strukturo ζ_α. Če je ciklična struktura ξ dopustna, tj. $\xi \xrightarrow{k} \zeta_\alpha$, potem obstaja rešitev enačbe φ s ciklično strukturo ξ.*

Dokaz. Dokaza ne bomo naredili v vsej splošnosti, temveč samo na konkretnem zgledu. Rešujmo enačbo

$$\varphi^2 = (1\,2)(3\,4)(5\,6\,7\,8\,9) = \alpha. \tag{8.10}$$

Dopustno ciklično strukturo uganemo, velja namreč $[4,5] \xrightarrow{2} [2,2,5]$, zato bomo permutacijo φ zapisali z nedoločenimi koeficienti,

$$\varphi = (a\,b\,c\,d)(e\,f\,g\,h\,i),$$

in izračunali njen kvadrat,

$$\varphi^2 = (a\,c)(b\,d)(e\,g\,i\,f\,h). \tag{8.11}$$

Če izenačimo istoležne koeficiente permutacij φ^2 in α iz enačb (8.10) in (8.11), potem lahko zapišemo rešitev φ takole

$$\varphi = (1\,3\,2\,4)(5\,8\,6\,9\,7).$$

□

Trditev 8.12 in uporaba metode nedoločenih koeficientov bo naša glavna strategija pri

reševanju potenčne enačbe (8.9). Določili bomo ciklično strukturo ζ_α permutacije α, nato pa poiskali *vse* dopustne ciklične strukture ξ, ki zadoščajo $\xi \xrightarrow{k} \zeta_\alpha$. Pri vsaki dopustni ciklični strukturi bomo z metodo nedoločenih koeficientov uspeli poiskati rešitev.

Na tem mestu omenimo, da opisana metoda nedoločenih koeficientov ne bo uspela poiskati *vseh* rešitev.

Analizo bomo stopnjevali v treh korakih.

Permutacija α je m-cikel

Po trditvi 8.4 lahko permutacijo α (ki ima ciklično strukturo [m]) zapišemo kot potenco permutacije φ samo v primeru, ko je ciklična struktura permutacije φ prav tako enaka $[m]$. Z drugimi besedami, $\xi = [m]$ je *edina* ciklična struktura, za katero je $\xi \to [m]$. Če velja celo $[m] \xrightarrow{k} [m]$, potem je $\gcd(m, k) = 1$. Ali obratno, če je α m-cikel in $\gcd(m, k) \geq 2$, potem enačba $\varphi^k = \alpha$ *ni* rešljiva.

Če pa je $\gcd(m, k) = 1$, potem je rešitev celo enolična in lahko jo poiščemo z znanjem o linearnih diofantskih enačbah. Linearna diofantska enačba

$$kx = my + 1$$

je namreč rešljiva, četudi je nismo zapisali v standardni obliki. Izberemo lahko celo rešitev, v kateri sta tako x kot y pozitivni števili. Zato je

$$\varphi = \mathrm{id} * \varphi = (\varphi^m)^y * \varphi = \varphi^{my+1} = (\varphi^k)^x = \alpha^x$$

edina rešitev enačbe (8.9).

V permutaciji α nastopajo cikli samih enakih dolžin

Denimo, da je ciklična struktura permutacije α enaka $[3, 3, 3, 3, 3]$. Katere so tiste ciklične strukture, za katere velja $\xi \to [3, 3, 3, 3, 3]$?

Cikel permutacije lahko pri potenciranju razpade na disjunktne 3-cikle samo takrat, ko je dolžina prvotnega cikla večkratnik dolžine 3. Tako lahko disjunktni 3-cikli nastanejo pri potenciranju 3-, 6-, 9-, 12- in 15-cikla (daljši cikli ne pridejo v poštev, saj bi dobili preveč 3-ciklov). Zanimivo je, da $[9, 3, 3] \not\to [3, 3, 3, 3, 3]$ in $[9, 6] \not\to [3, 3, 3, 3, 3]$. Če namreč 9-cikel razpade na tri 3-cikle, potem je eksponent deljiv s 3. V tem primeru pa 6- oziroma 3-cikel ne razpadeta na željen način.

Preostale možnosti so ugodne, velja namreč

$$[15] \to [3, 3, 3, 3, 3], \qquad [12, 3] \to [3, 3, 3, 3, 3], \qquad [6, 6, 3] \to [3, 3, 3, 3, 3],$$
$$[6, 3, 3, 3] \to [3, 3, 3, 3, 3], \quad [3, 3, 3, 3, 3] \to [3, 3, 3, 3, 3].$$

Kot zgled izberimo eksponent $k = 2$. V tem primeru so dopustne ciklične strukture $[3, 3, 3, 3, 3]$, $[6, 3, 3, 3]$ in $[6, 6, 3]$, saj velja $[6] \xrightarrow{2} [3, 3]$. Cikel dolžine 15 se pri kvadriranju ohrani, 10-cikel pa razpade na dva 5-cikla. Ustrezne rešitve, če poznamo še α, dobimo z metodo nedoločenih koeficientov.

V permutaciji α nastopajo disjunktni cikli različnih dolžin

V tem primeru nalogo razcepimo na več manjših nalog prejšnjega tipa. Permutacijo α pišemo kot produkt permutacij

$$\alpha = \alpha_{d_1} * \alpha_{d_2} * \cdots * \alpha_{d_\ell},$$

kjer z α_{d_i} označimo produkt vseh ciklov dolžine d_i iz zapisa permutacije α.

S φ_{d_i} označimo morebitno rešitev enačbe

$$\varphi_{d_i}^k = \alpha_{d_i}. \tag{8.12}$$

Če obstaja rešitev enačbe (8.12) za vse $i = 1, \ldots, \ell$, potem je

$$\varphi = \varphi_{d_1} * \varphi_{d_2} * \cdots * \varphi_{d_\ell}$$

rešitev enačbe (8.9). Če pa pri kakem $i \in \{1, \ldots, \ell\}$ enačba (8.12) ni rešljiva, potem tudi enačba (8.9) nima rešitve.

Kot zgled poiščimo rešitev enačbe

$$\varphi^k = (1\,2\,3)(4\,5\,6)(7\,8\,9)(10\,11)(12\,13)(14) \tag{8.13}$$

pri eksponentih $k = 543$ in $k = 544$. Enačbo razcepimo na tri enačbe

$$\begin{aligned}
\varphi_3^k &= (1\,2\,3)(4\,5\,6)(7\,8\,9),\\
\varphi_2^k &= (10\,11)(12\,13),\\
\varphi_1^k &= (14),
\end{aligned}$$

in opazujemo možne ciklične strukture permutacij φ_1, φ_2 in φ_3. Za φ_3 velja $[3,3,3] \to [3,3,3]$, $[6,3] \to [3,3,3]$ in $[9] \to [3,3,3]$. Za φ_2 ugotovimo $[2,2] \to [2,2]$ in $[4] \to [2,2]$, medtem ko je $[1]$ edina ciklična struktura, ki pri potenciranju pridela ciklično strukturo $[1]$.

Obdelajmo najprej prvi eksponent $k = 543$. Ker je $\gcd(543, 9) = 3$ in $\gcd(543, 3) = 3$, je za φ_3 dopustna samo ciklična struktura $[9]$. Permutacijo φ_3 zapišimo z nedoločenimi koeficienti,

$$\varphi_3 = (x_1\,x_2\,x_3\,x_4\,x_5\,x_6\,x_7\,x_8\,x_9)$$

in izračunajmo

$${\varphi_3}^{543} = {\varphi_3}^{543 \bmod 9} = {\varphi_3}^3 = (x_1\,x_4\,x_7)(x_2\,x_5\,x_8)(x_3\,x_6\,x_9).$$

Odtod preberemo možno delno rešitev $\varphi_3 = (1\,4\,7\,2\,5\,8\,3\,6\,9)$.

Ker je $\gcd(543, 2) = 1$, je edina dopustna ciklična struktura za φ_2 enaka $[2,2]$. Pišemo $\varphi_2 = (y_1\,y_2)(y_3\,y_4)$ in izračunamo njegovo potenco

$${\varphi_2}^{543} = {\varphi_2}^{543 \bmod 2} = {\varphi_2}^1 = (y_1\,y_2)(y_3\,y_4).$$

Zato je $\varphi_2 = (10\,11)(12\,13)$.

Produkt permutacij φ_1, φ_2 in φ_3 je permutacija

$$\varphi = (1\,4\,7\,2\,5\,8\,3\,6\,9)(10\,11)(12\,13)(14),$$

ki je rešitev enačbe (8.13) v primeru $k = 543$.

Če pa je eksponent $k = 544$, potem ne obstaja dopustna ciklična struktura za φ_2. Tako 4-cikel kot 2-cikel pri potenciranju na 544-to potenco razpadeta na cikle dolžine 1. Torej enačba (8.13) nima rešitve v primeru $k = 544$.

Poglavje 9

Grafi

9.1 Osnove

Graf je urejen par $G = (V, E)$, kjer je

- V neprazna končna množica *točk* grafa G in
- E množica *povezav* grafa G, pri čemer je vsaka povezava *par* točk[1].

Točki u, v grafa $G = (V, E)$ sta *sosedi*, če je $\{u, v\}$ povezava v grafu G. V tem primeru pravimo tudi, da sta u in v *krajišči* povezave $e = \{u, v\}$ in da se povezava e *dotika* svojih krajišč, v tem primeru točk u in v. Za par sosednjih točk u, v v grafu G uporabljamo tudi oznako $u \sim_G v$ (ali celo $u \sim v$, če graf G preberemo iz konteksta).

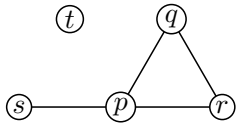
Povezavo $\{u, v\}$ bomo tipično zapisovali kot uv ali vu. Množico točk grafa G označimo z $V(G)$, množico povezav pa z $E(G)$. Za točko v z oznako $N(v)$ označimo množico sosed točke v, z oznako $E(v)$ pa množico vseh povezav, ki se dotikajo točke v.

Stopnja točke v v grafu G je enaka številu povezav $e \in E(G)$, ki se dotikajo točke v. Hkrati je enaka tudi številu sosed točke v. Stopnjo točke v označimo z $\deg(v)$. Točko stopnje 0 imenujemo tudi *izolirana* točka, točki stopnje 1 pravimo tudi *list*.

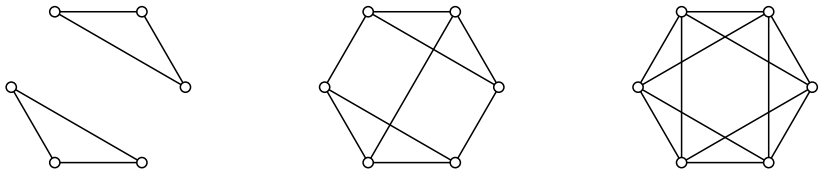
Oglejmo si graf $G = (\{p, q, r, s, t\}, \{pq, pr, qr, ps\})$. Graf G ima pet točk in štiri povezave in je predstavljen na sliki 9.1. Stopnje točk v grafu G so po vrsti enake $\deg(p) = 3$, $\deg(q) = \deg(r) = 2$, $\deg(s) = 1$ in $\deg(t) = 0$. Točka s je list in t je izolirana točka v grafu G.

Točke grafa narišemo kot krožce v ravnini, vsako povezavo pa kot krivuljo, tipično kar daljico, med risbama njenih krajišč.

[1] Spomnimo se, par je množica z natanko dvema elementoma. Za množici točk in povezav zahtevamo tudi, da sta disjunktni. Če je za graf G njegova množica točk enaka $\{x, y, \{x, y\}\}$ in sta točki x in y sosedi v G, potem je par $\{x, y\}$ po eni strani povezava grafa G in po drugi točka v grafu G. S tem formalno ni nič narobe. Povzroča pa popolnoma nepotrebno zmedo, ki se ji bomo na daleč izognili.



Slika 9.1: Risba grafa $G = (\{p, q, r, s, t\}, \{pq, pr, qr, ps\})$.



Slika 9.2: 2-regularen, kubičen (3-regularen) in 4-regularen graf.

Graf G je *regularen* graf, če imajo vse njegove točke isto stopnjo. Natančneje, graf G je *d-regularen*, če so vse točke stopnje d; 3-regularnim grafom pravimo tudi *kubični* grafi. V grafu G z $\Delta(G)$ označimo največjo in z $\delta(G)$ najmanjšo stopnjo katere od njegovih točk. Regularni grafi so natanko tisti grafi, za katere je $\Delta(G) = \delta(G)$. Primeri regularnih grafov so predstavljeni na sliki 9.2.

Vsota stopenj točk grafa je enaka dvakratniku števila povezav. Simbolično:

Trditev 9.1 *Naj bo G graf. Potem je*

$$\sum_{v \in V(G)} \deg(v) = 2 \cdot |E(G)| \tag{9.1}$$

Dokaz. Opazujmo moč množice urejenih parov

$$\mathcal{F} = \{(v, e) \mid \text{točka } v \text{ je krajišče povezave } e\} \subseteq V(G) \times E(G).$$

Število takšnih parov je po eni strani enako $2 \cdot |E(G)|$, saj je vsaka povezava $e \in E(G)$ druga koordinata v natanko dveh takšnih parih — po enem za vsako krajišče.

Po drugi strani je prispevek vsake točke $v \in V(G)$ k moči $\mathcal{F}$ enak njeni stopnji. Točka v je namreč prva koordinata v natanko $\deg(v)$ takšnih parih — po enem za vsako povezavo e, ki se dotika točke v. Torej je $|\mathcal{F}| = \sum_{v \in V(G)} \deg(v)$. □

Trditev 9.1 v družini regularnih grafov pomeni tesno zvezo med številom točk in povezav.

Trditev 9.2 *Če je G d-regularen graf z n točkami in m povezavami, potem velja*

$$d \cdot n = 2 \cdot m.$$

Druga posledica trdi, da je v vsakem grafu sodo mnogo točk lihe stopnje.

Trditev 9.3 *Vsak graf G vsebuje sodo mnogo točk lihe stopnje.*

Dokaz. Vsoto stopenj točk razpišemo kot vsoto dveh delnih vsot. V prvi delni vsoti seštejemo stopnje točk sodih stopenj, v drugi stopnje točk lihih stopenj.

$$\sum_{v \in V(G)} \deg(v) = \sum_{\substack{v \in V(G) \\ v \text{ sode stopnje}}} \deg(v) + \sum_{\substack{v \in V(G) \\ v \text{ lihe stopnje}}} \deg(v)$$

Delni vsoti sta iste parnosti, saj je njuna vsota po trditvi 9.1 soda. Ker je vsota sodih stopenj soda, mora biti takšna tudi vsota lihih stopenj. To je možno samo v primeru, ko je število členov v tej vsoti sodo. Zato ima G sodo mnogo točk lihe stopnje. □

Osnovne operacije z grafi

Oglejmo si nekaj osnovnih operacij za delo z grafi.

Naj bo $G = (V, E)$ graf ter u in v njegovi točki.

Če sta točki u in v sosedi v grafu G, $uv \in E(G)$, potem lahko grafu G povezavo uv *odstranimo*. Dobljeni graf označimo z $G - uv$ in zanj velja

$$V(G - uv) = V(G) \quad \text{in} \quad E(G - uv) = E(G) \setminus \{uv\}.$$

Če točki u in v v grafu G nista sosedi, $uv \notin E(G)$, potem lahko grafu G povezavo uv *dodamo*. Dobljeni graf označimo z $G + uv$ in zanj velja

$$V(G + uv) = V(G) \quad \text{in} \quad E(G + uv) = E(G) \cup \{uv\}.$$

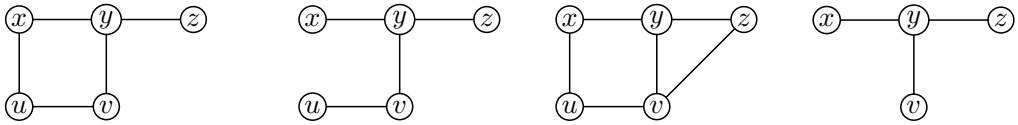
Zapis odstranjevanja in dodajanja povezav lahko interpretiramo tudi v razširjenem smislu. Četudi e ni povezava v grafu G, smemo pisati $G - e$. V tem primeru je graf $G - e$ seveda enak originalnemu grafu G. Ravno tako smemo zapisati $G + e'$, čeprav je $e' \in E(G)$. Tudi v tem primeru velja zveza $G + e' = G$.

Operaciji odstranjevanja in dodajanja povezav sta predstavljeni na sliki 9.3.

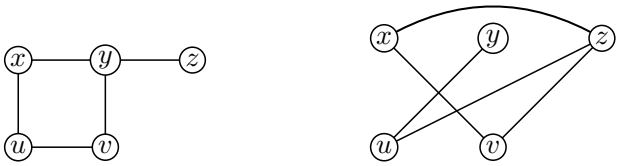
Operacije s točkami zahtevajo nekaj pazljivosti. Če grafu G odstranimo točko v, potem je potrebno hkrati odstraniti tudi vse povezave iz $E(v)$. Zato je $G - v$ graf, ki zadošča

$$V(G - v) = V(G) \setminus \{v\} \quad \text{in} \quad E(G - v) = E(G) \setminus E(v).$$

Dodajanja nove točke grafu G ne moremo odpraviti tako enostavno. Potrebno je namreč opisati tudi, katere *nove* povezave se dotikajo *sveže dodane* točke.



Slika 9.3: Graf G ter grafi $G - ux$, $G + vz$ in $G - u$.



Slika 9.4: Graf G in njegov komplement $\overline{G}$.

Operaciji odstranjevanja točk in povezav lahko posplošimo na množice točk in povezav. Ravno tako lahko dodamo celo množico povezav.

Naj bo $G = (V, E)$ graf, $U \subseteq V(G)$ podmnožica njegovih točk in F množica parov točk iz $V(G)$. Z $G - U$ označimo graf, ki ga dobimo z zaporednim odstranjevanjem točk iz U. Podobno z $G - F$ označimo graf, ki ga dobimo z zaporednim odstranjevanjem povezav iz F; $G - F$ je graf z množico točk $V(G)$ in množico povezav $E(G) \setminus F$.

Z $G + F$ označimo graf, ki ga dobimo z zaporednim dodajanjem povezav iz F; $G + F$ je graf z množico točk $V(G)$ in množico povezav $E(G) \cup F$.

Definirati smemo tudi operaciji preseka in unije grafov. Za grafa G in H je njun *presek* graf $G \cap H$, za katerega je

$$V(G \cap H) = V(G) \cap V(H) \quad \text{in} \quad E(G \cap H) = E(G) \cap E(H). \tag{9.2}$$

Podobno je za grafa G in H njuna *unija* graf $G \cup H$, za katerega velja

$$V(G \cup H) = V(G) \cup V(H) \quad \text{in} \quad E(G \cup H) = E(G) \cup E(H). \tag{9.3}$$

Če poleg tega velja tudi $V(G) \cap V(H) = \emptyset$, pravimo, da je $G \cup H$ *disjunktna unija* grafov G in H.

Definirajmo še eno grafovsko operacijo. *Komplement* grafa G, označimo ga z $\overline{G}$, je graf z isto množico točk kot G, pri čemer sta različni točki u in v sosedi v $\overline{G}$ natanko tedaj, ko nista sosedi v G. Velja torej

$$V(\overline{G}) = V(G) \quad \text{in} \quad E(\overline{G}) = \{uv \mid u \neq v \text{ in } uv \notin E(G)\}.$$

Graf G in njegov komplement $\overline{G}$ sta predstavljena na sliki 9.4.

Drugačni razredi grafov

V našem modelu grafov smo se omejili na grafe s končno mnogo točkami, pri čemer sta točki v grafu lahko sosedi ali pa ne.

Zakaj graf ne bi smel imeti *neskončne* množice točk? V tem primeru ima posamezna točka lahko celo neskončno mnogo sosed, vsekakor pa bi težko govorili o vsoti stopenj točk.

V *usmerjenih grafih* imajo povezave usmeritve, vsaka povezava ima začetno in končno točko. V tem primeru je možno povezave opisati z urejenimi pari točk. Za posamezno točko v govorimo o njeni izhodni oziroma vhodni stopnji (številu povezav, ki imajo točko v za začetek oziroma konec).

Multigraf je graf, v katerem dovolimo vzporedne povezave — obstaja lahko več povezav, ki imajo vse isti par točk za krajišči. Povezave moramo v tem primeru definirati precej drugače kot s pari točk. Potrebna je večja pazljivost pri definiciji stopnje točke, saj se število sosed točke lahko razlikuje od števila povezav, ki se posamezne točke dotikajo.

Podobno velja za *zanke*, povezave, ki imajo obe krajišči v isti točki. Kakšen je prispevek zanke k stopnji posamezne točke?

Takšne razširitve domene so včasih dobrodošle in naravne. V praksi je še največ težav z neskončnimi grafi (če nič drugega, jih težko predstavimo v računalniku), medtem ko usmerjeni grafi ali grafi z vzporednimi povezavami ne povzročajo tehničnih preglavic.

Vseeno se omejimo na neusmerjene grafe brez vzporednih povezav, ki vsebujejo končno mnogo točk. Že na tem (pohlevnem) grafovskem razredu bomo spoznali zadosten nabor konstrukcij in grafovskih vprašanj, da usmerjenih in vzporednih povezav niti ne bomo pogrešali.

9.2 Izomorfizem grafov

Grafa G_1 in G_2 sta predstavljena na sliki 9.5. Sta morda enaka ali različna?

Slika 9.5: Izomorfna grafa G_1 in G_2.

Seveda sta različna. Graf je definiran kot urejen par množic točk in povezav, enakost urejenih parov pomeni ujemanje po posameznih komponentah. Grafa G_1 in G_2 imata

sicer isti množici točk, toda točki 3 in 4 sta v grafu G_1 sosedi, medtem ko v grafu G_2 nista.

Po drugi strani pa, če zanemarimo oznake točk, lahko z rotacijo enega od grafov za 180° dosežemo ujemanje njunih risb. Torej grafa G_1 in G_2 nista močno različna.

V tem razdelku bomo izdelali formalen način obravnave grafov, ki se razlikujejo samo v oznakah točk.

Grafa $G = (V, E)$ in $G' = (V', E')$ sta *izomorfna*, če obstaja preslikava

$$\varphi : V(G) \to V(G'), \tag{9.4}$$

ki zadošča:

(IZO1) φ je bijektivna preslikava iz $V(G)$ v $V(G')$ in

(IZO2) za vsaki točki $u, v \in V(G)$ velja $uv \in E(G) \iff \varphi(u)\varphi(v) \in E(G')$.

Preslikavo φ (9.4) v tem primeru imenujemo *izomorfizem* med grafoma G in G' in pišemo $G \cong G'$.

Ustrezno preslikavo $\varphi_{12} : V(G_1) \to V(G_2)$, izomorfizem med grafoma G_1 in G_2 s slike 9.5, lahko predstavimo kar s tabelo

x	1	2	3	4
$\varphi_{12}(x)$	3	4	1	2

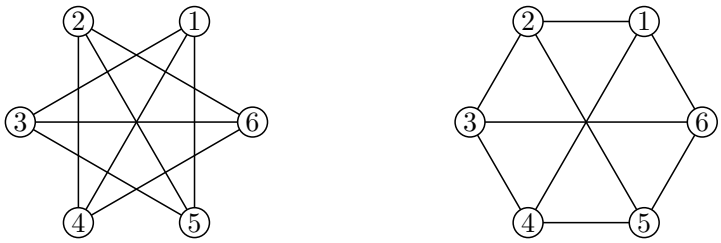
(9.5)

Pod drobnogled vzemimo naslednji par grafov, grafa H_1 in H_2, ki sta predstavljena na sliki 9.6. Na prvi pogled se zdi, da nista izomorfna. Lahko to utemeljimo?

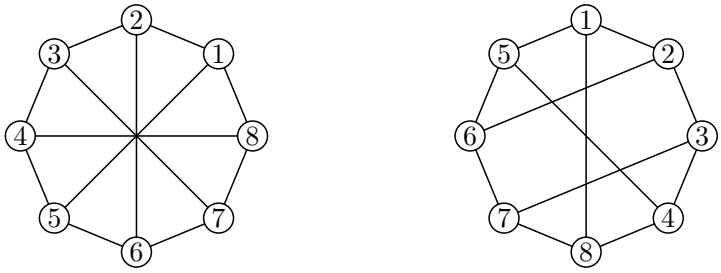
Slika 9.6: Neizomorfna grafa H_1 in H_2.

Grafa H_1 in H_2 imata isto število točk, kar je seveda potreben pogoj. Če naj bo preslikava ψ_{12} domnevni izomorfizem med grafoma H_1 in H_2, potem mora veljati $|V(H_1)| = |V(H_2)|$, saj mora biti ψ_{12} bijektivna preslikava iz $V(H_1)$ v $V(H_2)$.

Opazujmo točko $2 \in V(H_1)$, ki je soseda z vsemi preostalimi točkami grafa H_1. Slika točke 2, $\psi_{12}(2)$, mora biti po (IZO2) soseda z vsemi točkami $\psi_{12}(1)$, $\psi_{12}(3)$ in $\psi_{12}(4)$, ki so po (IZO1) *različne* točke grafa H_2. To pa ni možno, saj nobena izmed točk grafa H_2 ni stopnje 3.



Slika 9.7: Neizomorfna grafa H_3 in H_4.



Slika 9.8: Izomorfna grafa G_3 in G_4.

Grafa H_1 in H_2 nista izomorfna, $H_1 \ncong H_2$. Imata sicer isto število točk in tudi isto število povezav. Vseeno nista izomorfna, izomorfizem med H_1 in H_2 ne obstaja. Graf H_1 namreč vsebuje točko stopnje 3, graf H_2 takšne točke nima.

V naslednjem zgledu opazujmo grafa H_3 in H_4 s slike 9.7. Znova se vprašajmo, ali sta izomorfna. Morebitni izomorfizem med H_3 in H_4 označimo s črko ψ_{34}.

Grafa H_3 in H_4 imata isto število točk in povezav ter celo isto zaporedje stopenj točk — oba sta namreč kubična grafa, vse njune točke so stopnje 3.

Graf H_3 vsebuje *trikotnik*, paroma sosednje točke 1, 3 in 5. Izomorfizem $\psi_{34} : V(H_3) \to V(H_4)$, če seveda obstaja, mora zagotoviti, da bodo $\psi_{34}(1)$, $\psi_{34}(3)$ in $\psi_{34}(5)$ paroma sosednje točke v grafu H_4. Toda graf H_4 ne vsebuje trojice paroma sosednjih točk. Posledično izomorfizem ψ_{34} med grafoma H_3 in H_4 ne obstaja. Grafa H_3 in H_4 nista izomorfna, $H_3 \ncong H_4$.

Kot zadnji zgled opazujmo grafa G_3 in G_4, predstavljena sta na sliki 9.8. Grafa sta izomorfna, saj imata natanko iste povezave. Preslikava iz $V(G_3)$ v $V(G_4)$, iskani izomorfizem, je kar identiteta[2]. Velja torej $G_1 \cong G_2$.

V nadaljevanju pokažemo, da izomorfizem porodi ekvivalenčno relacijo.

Izrek 9.4 *Relacija izomorfnosti grafov $\cong$ je ekvivalenčna relacija v množici grafov.*

Dokaz. Za poljuben graf G velja $G \cong G$, ustrezna preslikava je kar identiteta.

[2]Za grafa G_3 in G_4 bi lahko rekli, da gre za *isti* graf, ki je predstavljen z dvema različnima risbama.

Če je $G_1 \cong G_2$, potem obstaja izomorfizem $\varphi : V(G_1) \to V(G_2)$. Inverzna preslikava $\varphi^{-1} : V(G_2) \to V(G_1)$ je po izreku 5.5 ravno tako bijektivna preslikava. Izberimo poljubni točki $v_2, u_2 \in V(G_2)$. Ker je preslikava φ bijekcija, v grafu G_1 obstajata enolično določeni točki v_1, u_1, za kateri je $\varphi(v_1) = v_2$ in $\varphi(u_1) = u_2$ oziroma $\varphi^{-1}(v_2) = v_1$ in $\varphi^{-1}(u_2) = u_1$. Ker je φ izomorfizem, sta točki v_2, u_2 sosedi v grafu G_2 natanko tedaj, ko sta v grafu G_1 sosedi točki $v_1 = \varphi^{-1}(v_2)$ in $u_1 = \varphi^{-1}(u_2)$. Torej je preslikava φ^{-1} res izomorfizem.

Za tranzitivnost izomorfizma je dovolj premisliti, da je kompozitum izomorfizmov tudi izomorfizem. Kompozitum bijektivnih preslikav je po izreku 5.7 bijektivna preslikava, ravno tako pa kompozitum izomorfizmov ohranja lastnost sosednosti oziroma nesosednosti parov točk. □

Razdelek končajmo z razlago grafovske lastnosti. Če graf G "*vsebuje točko stopnje* 2", potem jo vsebuje tudi vsak graf G', ki je izomorfen grafu G. Ravno tako izomorfna grafa G in G' hkrati zadoščata vsaki od lastnosti "*ima natanko* 32 *povezav*", "*vsebuje trojico paroma sosednjih točk*" oziroma "*vsebuje sosednji točki, ki sta stopenj* 2 *in* 3."

Grafovska lastnost je, formalno, podmnožica grafov, ki ji hkrati z grafom G pripadajo tudi vsi grafi G', ki so izomorfni G. Lastnost "*ima natanko* 32 *povezav*" si (formalno) predstavljamo kot množico vseh grafov z natanko 32 povezavami.

Množica vseh grafov, v katerih "*sta točki u in v sosedi*", ni grafovska lastnost. Pri grafu G s povezavo uv se je moč s preimenovanjem točke u v u' znebiti povezave uv.

Kako torej utemeljimo odgovor na vprašanje "*Ali sta grafa G in G′ izomorfna?*"

Pozitivni odgovor najlaže utemeljimo s konstrukcijo izomorfizma med G in G'. Za negativni odgovor pa je dovolj poiskati grafovsko lastnost, ki ji zadošča natanko eden od grafov G, G'.

9.3 Osnovne družine grafov

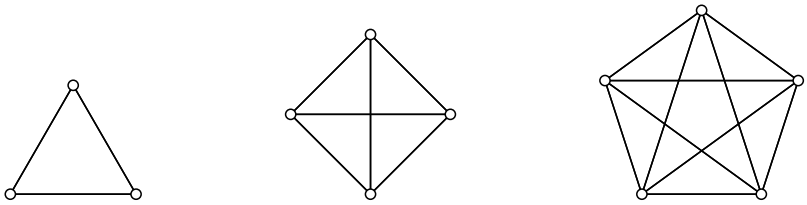
Pod streho smo spravili razdelek o izomorfizmu grafov. Izrek 9.4 trdi, da je izomorfnost grafov ekvivalenčna relacija v množici grafov. V tem razdelku definiramo nekatere standardne družine grafov (natančneje, ekvivalenčne razrede grafov za izomorfizem) — predstavnike teh ekvivalenčnih razredov grafov pa opišemo s konkretnimi množicami točk in povezav.

Polni grafi

Graf G je *poln*, če sta vsaki dve njegovi točki sosedi. Polni graf na $n \geq 1$ točkah označimo[3] s K_n. Množici točk in povezav polnega grafa K_n *lahko* opišemo takole:

$$V(K_n) = \{v_1, \ldots, v_n\},$$
$$E(K_n) = \{v_i v_j \mid 1 \leq i < j \leq n\}.$$

Polni graf K_n ima, po definiciji, natanko n točk in $\frac{n(n-1)}{2}$ povezav in je $(n-1)$-regularen graf. Polni grafi K_3, K_4 in K_5 so predstavljeni na sliki 9.9.



Slika 9.9: Polni grafi K_3, K_4 in K_5.

Prazni grafi

Graf G brez povezav imenujemo *prazen*. Prazni graf na $n \geq 1$ točkah označimo s $\overline{K_n}$. Množici točk in povezav praznega grafa $\overline{K_n}$ lahko opišemo takole:

$$V(\overline{K_n}) = \{v_1, \ldots, v_n\},$$
$$E(\overline{K_n}) = \emptyset.$$

Prazni graf $\overline{K_n}$ vsebuje n točk in 0 povezav. Prazni graf $\overline{K_n}$ je komplement polnega grafa K_n, pri $n = 1$ pa velja celo zveza[4] $K_1 = \overline{K_1}$.

Polni dvodelni grafi

Graf G je *poln dvodelen* graf, če lahko njegove točke razbijemo v dva bloka (tudi *barvna razreda*), pri čemer sta točki sosedi natanko tedaj, ko pripadata različnima blokoma. S $K_{m,n}$, pri čemer je $m \geq 1$ ali $n \geq 1$, označimo polni dvodelni graf na $m + n$ točkah, ki

[3] Oznaka K_7 dejansko označuje ekvivalenčni razred grafov (za relacijo izomorfnosti), ki imajo natanko 7 točk in v katerih sta vsaki dve točki sosedi, in ne zgolj posameznega grafa iz tega ekvivalenčnega razreda.

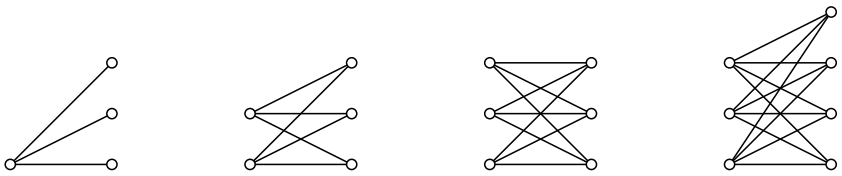
[4] Primerjamo ekvivalenčna razreda grafov K_1 in $\overline{K_1}$, ki sta seveda enaka. Konkretna predstavnika razredov K_1 in $\overline{K_1}$ sta morda zgolj izomorfna.

ima v enem bloku m in v drugem n točk. Točke in povezave grafa $K_{m,n}$ lahko opišemo na naslednji način:

$$V(K_{m,n}) = \{u_1, \ldots, u_m, v_1, \ldots, v_n\},$$
$$E(K_{m,n}) = \{u_i v_j \mid 1 \leq i \leq m \text{ in } 1 \leq j \leq n\}.$$

Graf $K_{m,n}$ vsebuje $m + n$ točk in $m \cdot n$ povezav. Točke grafa $K_{m,n}$ so stopenj m ali n, graf $K_{m,n}$ je regularen samo v primeru, ko je $m = n$.

Pri tem velja enakost $K_{1,1} = K_2$. Nekaj zgledov polnih dvodelnih grafov je predstavljenih na sliki 9.10.



Slika 9.10: Polni dvodelni grafi $K_{1,3}$, $K_{2,3}$, $K_{3,3}$ in $K_{3,4}$.

Cikli

Cikel na $n \geq 3$ točkah označimo s C_n. Množica točk cikla C_n je enaka množici ostankov po modulu n, točki-ostanka pa sta sosedi natanko tedaj, ko se razlikujeta za 1. Pišemo lahko:

$$V(C_n) = Z_n = \{0, \ldots, n-1\},$$
$$E(C_n) = \{uv \mid v \equiv u + 1 \pmod{n}\}.$$

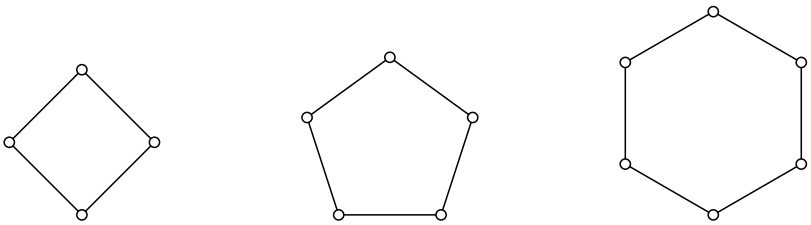
Ciklu z n točkami rečemo tudi *n-cikel*. Ciklom s sodo (liho) mnogo točkami pravimo tudi *sodi (lihi)* cikli.

Cikel C_n vsebuje natanko n točk in n povezav, vse točke cikla C_n so stopnje 2. Nekaj ciklov je predstavljenih na sliki 9.11. Velja tudi $C_3 = K_3$ in $C_4 = K_{2,2}$.

Poti

Pot na n točkah označimo s P_n. Za točke poti P_n lahko izberemo naravna števila iz množice $[n]$, točki pa sta sosedi natanko tedaj, ko se razlikujeta za 1 (za razliko od ciklov se pri $n \geq 3$ naravni števili 0 in $n-1$, najmanjši in največji element množice $[n]$ (in ne ostanka pri deljenju z n), ne razlikujeta zgolj za 1).

$$V(P_n) = [n] = \{0, \ldots, n-1\},$$
$$E(P_n) = \{uv \mid v = u + 1\}.$$



Slika 9.11: Cikli C_4, C_5 in C_6.

Slika 9.12: Poti P_4, P_5 in P_6.

Pri $n \geq 2$ pot P_n vsebuje natanko dve točki stopnje 1, ki jima pravimo tudi *krajišči*. Preostale točke so stopnje 2 in jih imenujemo *notranje točke*.

Veljajo zveze $P_1 = K_1$, $P_2 = K_2$ in $P_3 = K_{1,2}$. Pot P_n, $n \geq 3$, dobimo iz cikla C_n tako, da slednjemu odstranimo eno povezavo[5]. Nekaj zgledov poti je predstavljenih na sliki 9.12.

Hiperkocke

Izberimo naravno število n; *n-razsežna hiperkocka* je graf, ki ga označimo s Q_n. Njene točke so dvojiška zaporedja (členi so enaki 0 oziroma 1) dolžine n, dve točki-zaporedji pa sta sosedi, če se razlikujeta v natanko enem členu. Formalno zapišemo takole:

$$V(Q_n) = \{b_1 \dots b_n \mid b_i \in \{0,1\}, i \in \{1,\dots,n\}\},$$
$$E(Q_n) = \{(b_1 \dots b_n)\,(b'_1 \dots b'_n) \mid \exists i \in \{1,\dots,n\} \forall j \in \{1,\dots,n\} (b_j = b'_j \Leftrightarrow i \neq j)\}.$$

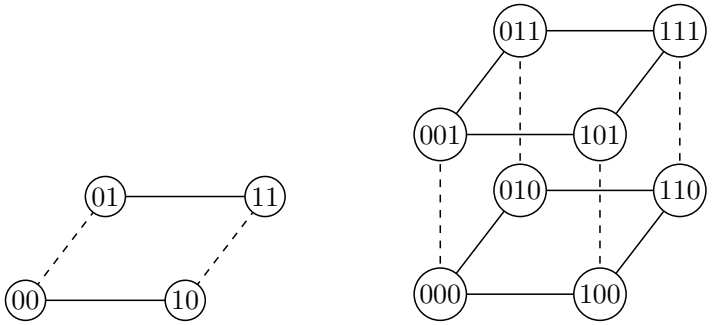
Dvojiško zaporedje dolžine $n \geq 1$ iz krajšega dvojiškega zaporedja ζ dolžine $n-1$ pridelamo tako, da zaporedju ζ na koncu pripišemo bodisi 0 bodisi 1. Zato lahko pri $n \geq 1$ hiperkocko Q_n konstruiramo iz dveh kopij $(n-1)$-razsežnih hiperkock, če dodamo vse povezave med istoležnimi točkami[6].

Na slikah 9.13 in 9.14 so predstavljene hiperkocke[7] Q_2, Q_3 in Q_4. Ravno tako je s črtkanimi povezavami nakazano, kako posamezno hiperkocko, denimo Q_3, pridelamo iz dveh kopij manjše hiperkocke, v tem primeru Q_2, z dodajanjem povezav med *istoležnimi* točkami v obeh kopijah.

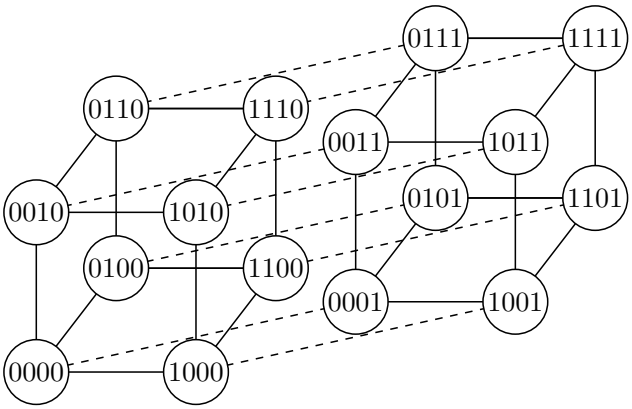
[5] Zelo pedantno bi se izrazili takole: Če je graf G cikel na $n \geq 3$ točkah in $e \in E(G)$, potem je $G - e$ pot na n točkah.

[6] Izberimo dve kopiji $(n-1)$-razsežne hiperkocke, označimo ju s Q in Q'. Točkam v Q, gre za dvojiška zaporedja dolžine $n-1$, na koncu pripišimo 0, točkam v Q' pa pripišimo 1. *Istoležni* točki iz Q in Q' sta tisti, ki se razlikujeta samo v zadnjem členu. Če dodamo vse povezave med pari istoležnih točk, pridelamo n-razsežno hiperkocko.

[7] Odkod ime? Hiperkocka Q_3 je skelet geometrijskega telesa kocke — oglišča so točke, oglišči na skupnem robu sta sosedi.



Slika 9.13: Hiperkocki Q_2 in Q_3.



Slika 9.14: Hiperkocka Q_4.

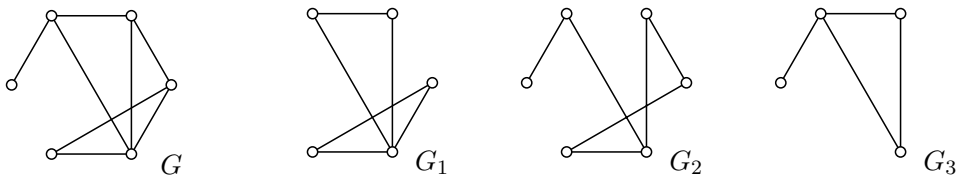
Hiperkocka Q_n ima natanko 2^n točk in je n-regularen graf — posamezna točka (dvojiško zaporedje dolžine n) je soseda s tistimi zaporedji, ki jih dobimo s spremembo natanko enega izmed n členov zaporedja. Po trditvi 9.2 izračunamo tudi število povezav hiperkocke, $|E(Q_n)| = n \cdot 2^{n-1}$.

Za hiperkocke veljajo naslednje zveze: $Q_1 = K_2$, $Q_2 = C_4$, pa tudi (prazno zaporedje je edino dvojiško zaporedje dolžine 0) $Q_0 = K_1$.

9.4 Podgrafi

Kaj je manjši in kaj je večji graf? Kako z grafovskimi operacijami definirati relacijo v množici grafov, ki je ustreznica relacije vsebovanosti pri množicah?

Naj bo $G = (V, E)$ graf in izberimo podmnožico točk $V' \subseteq V$ in podmnožico povezav $E' \subseteq E$. Urejeni par (V', E') morda ni graf, saj krajišči povezave $e \in E'$ ne pripadata



Slika 9.15: Graf G in podgrafi G_1, G_2, G_3; G_2 je vpet in G_3 induciran.

nujno množici V'. Naivna izbira podmnožice točk in podmnožice povezav prvotnega grafa morda ne pomeni hkrati tudi grafovske strukture.

Ustrezno podstrukturo je potrebno definirati z nekoliko več pazljivosti. Pravimo, da je *graf* $G_1 = (V_1, E_1)$ *podgraf* grafa $G = (V, E)$, kar označimo tudi z $G_1 \subseteq G$, če velja tako $V_1 \subseteq V$ kot tudi $E_1 \subseteq E$. Ali enakovredno, graf G_1 lahko iz grafa G dobimo z zaporednim odstranjevanjem točk in/ali povezav, če začnemo z grafom G.

Kaj pa če dovolimo zgolj odstranjevanje povezav? Ali pa samo odstranjevanje točk? V teh dveh primerih pridelamo relaciji vpetega in induciranega podgrafa.

Graf $G_2 = (V_2, E_2)$ je *vpet podgraf* grafa $G = (V, E)$, če je $G_2 \subseteq G$ in velja $V(G_2) = V(G)$. To simbolično zapišemo z $G_2 \subseteq_s G$. Vpeti podgrafi grafa G imajo vsi isto množico točk kot G, pridelamo jih lahko zgolj z odstranjevanjem povezav iz prvotnega grafa G.

Graf $G_3 = (V_3, E_3)$ je *induciran podgraf* grafa G, $G_3 \subseteq_i G$, če lahko graf G_3 pridelamo iz grafa G že z zaporednim odstranjevanjem točk. Če je G_3 inducirani podgraf grafa G, potem velja

$$\forall u, v \in V(G_3)(uv \in E(G) \Rightarrow uv \in E(G_3)). \tag{9.6}$$

Z drugimi besedami, za točki u, v iz induciranega podgrafa G_3 lahko njuno morebitno sosednost preberemo kar iz prvotnega grafa G.

Na sliki 9.15 so prikazani graf G in trije njegovi podgrafi G_1, G_2 in G_3. Pri tem je G_2 vpet podgraf grafa G in G_3 induciran podgraf grafa G.

Podgrafi, operacije in relacije

Relacije $\subseteq$, $\subseteq_s$ in $\subseteq_i$ so delne urejenosti v množici grafov. Refleksivnost vseh omenjenih relacij upravičimo tako, da za konstrukcijo podgrafa (tudi vpetega oziroma induciranega) dopustimo tudi odstranjevanje nič točk in nič povezav iz prvotnega grafa G. Tranzitivnost sledi iz opazke, da lahko z odstanjevanjem točk/povezav nadaljujemo, tudi če se vmes za trenutek ustavimo. Antisimetričnost sledi iz iste lastnosti vsebovanosti. Če namreč za grafa G in H velja $G \subseteq H$ in $H \subseteq G$, potem veljajo tudi zveze $V(G) \subseteq V(H)$, $V(H) \subseteq V(G)$, $E(G) \subseteq E(H)$ in $E(H) \subseteq E(G)$. Ker je vsebovanost množic antisimetrična relacija, velja $V(H) = V(G)$ in $E(G) = E(H)$. Podoben sklep velja tudi za relaciji vpetega in induciranega podgrafa.

V množici grafov definirajmo relacijo R_s z opisom:

$G' R_s G$ natanko tedaj, ko G' dobimo iz G z odstranjevanjem natanko ene povezave.

Relacija vpetega podgrafa je v tem primeru tranzitivno-refleksivna ovojnica relacije R_s — simbolično: $H \subseteq_s G$ natanko tedaj, ko je $H R_s^* G$.

Podobno lahko definiramo relacijo R_i z opisom:

$G' R_i G$ natanko tedaj, ko G' dobimo iz G z odstranjevanjem natanko ene točke.

Relacija induciranega podgrafa je tranzitivno-refleksivna ovojnica relacije R_i. Relacija podgraf pa je trazitivno-refleksivna ovojnica relacije $R_s \cup R_i$.

Za konec še oznaka. Inducirani podgraf grafa G je določen s podmnožico svojih točk. Če je $U \subseteq V(G)$, je $G[U]$ inducirani podgraf grafa G z množico točk enako U. Velja torej

$$G[U] = G - (V(G) \setminus U).$$

9.5 Sprehodi v grafih

V tem razdelku bomo formalno izdelali pojem grafovskega sprehoda. Povezavo uv s krajiščema u in v smemo dojemati kot cesto, ki povezuje bližnji lokaciji u in v v večji strukturi. Če pa točki u in w nista sosedi, potem direktnega koraka med u in w ne moremo narediti,

Sprehod S v grafu G je zaporedje točk

$$v_0 v_1 v_2 \ldots v_{k-1} v_k, \tag{9.7}$$

pri čemer sta vsaki zaporedni točki sprehoda S, v_i in v_{i+1}, $i \in \{0, \ldots, k-1\}$, sosedi v grafu G, $v_i v_{i+1} \in E(G)$. V splošnem ne zahtevamo, da so točke v zaporedju (9.7) paroma različne.

Točko v_0 imenujemo tudi *začetek*, točko v_k pa *konec* sprehoda S (9.7). Z eno besedo, točki v_0 in v_k sta *krajišči* omenjenega sprehoda.

Sprehodu z začetkom v točki u in koncem v točki v pravimo tudi *u-v-sprehod*.

Zaporedni točki v_i in v_{i+1}, $i \in \{0, \ldots, k-1\}$, sprehoda S (9.7) sta sosedi v grafu G, zato pravimo, da sprehod S *uporabi* ali tudi *prehodi* povezavo $v_i v_{i+1}$. Še natančneje, sprehod S uporabi povezavo $v_i v_{i+1}$ *v smeri od* v_i do v_{i+1}.

Sprehod sme isto povezavo uporabiti večkrat, ravno tako lahko sprehod posamezno povezavo uporabi v različnih smereh. *Dolžino* sprehoda S označimo z $|S|$ in je enaka številu uporabljenih povezav (pri čemer upoštevamo njihovo večkratnost). Dolžina sprehoda S (9.7) je enaka k.

S sprehodi definiramo nekaj operacij. Izberimo indeksa $i, j \in \{0, \ldots, k\}$, $i \leq j$. *Odsek* sprehoda S (9.7) med v_i in v_j je sprehod

$$v_i v_{i+1} \ldots v_{j-1} v_j.$$

Odsek sprehoda S med v_i in v_j označimo tudi z $S_{v_i - v_j}$.

Obratni sprehod sprehoda S (9.7), označimo ga z S^R, je zaporedje

$$v_k v_{k-1} \ldots v_2 v_1 v_0. \tag{9.8}$$

Sprehoda S in S^R sta iste dolžine in imata isti krajišči.

Za sprehoda S_1 in S_2,

$$S_1 = u_1 u_2 \ldots u_{i-1} u_i \quad \text{in} \quad S_2 = u_i u_{i+1} \ldots u_{k-1} u_k,$$

kjer je začetek sprehoda S_2 enak koncu sprehoda S_1, lahko definiramo njun *stik*, sprehod

$$S_1 S_2 = u_1 u_2 \ldots u_{i-1} u_i u_{i+1} \ldots u_{k-1} u_k. \tag{9.9}$$

Dolžina stika $S_1 S_2$ je enaka vsoti dolžin sprehodov S_1 in S_2, $|S_1 S_2| = |S_1| + |S_2|$.

Opišimo še nekaj posebnih vrst sprehodov. Sprehod S (9.7) je *obhod*, če velja $v_0 = v_k$. Sprehod je obhod natanko tedaj, ko začetna in končna točka sovpadata.

Če sprehod nobene povezave ne uporabi dvakrat (ali večkrat), mu pravimo tudi *enostaven sprehod*.

Sprehod S (9.7) je *pot*, če so njegove točke paroma različne. Za različna $i, j \in \{0, \ldots, k\}$ mora veljati $v_i \neq v_j$.

Obhodu dolžine vsaj tri, v katerem so vse točke paroma različne (z izjemo ujemanja začetka in konca), pravimo *cikel*. Cikel[8] je torej sprehod

$$w_0 w_1 \ldots w_{k-1} w_k,$$

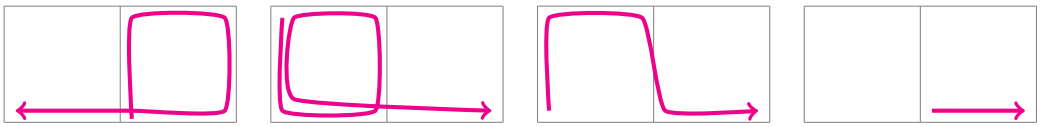
za katerega velja $k \geq 3$, $w_0 = w_k$ in za vsaka različna $i, j \in \{0, \ldots, k\}$ tudi implikacija $w_i = w_j \Rightarrow \{i, j\} = \{0, k\}$.

Poti in cikli so enostavni sprehodi.

Sprehode, poti, obhode in cikle v grafu najlaže prikažemo grafično, glej sliki 9.16 in 9.17. Omenimo še, da je zaporedje z enim samim členom v, če je v točka grafa G, sprehod (celo obhod) dolžine 0.

V tem razdelku navedimo še nekaj tehničnih rezultatov. Prvi govori o enakovrednosti obstoja sprehoda in poti med izbranima krajiščema.

[8] Termina *pot* in *cikel* imata dva različna pomena. Po eni strani je pot vrsta grafa, po drugi strani pa vrsta sprehoda, posebno zaporedje točk. Če je S pot v grafu G in iz grafa odstranimo vse povezave, ki jih S ne uporabi, in vse točke, ki niso v S, pridelamo graf, ki je izomorfen póti. Podobno velja za cikle, zmeda torej ne bo prehuda.



Slika 9.16: Prikazi sprehodov in poti.

Slika 9.17: Prikazi obhodov in ciklov.

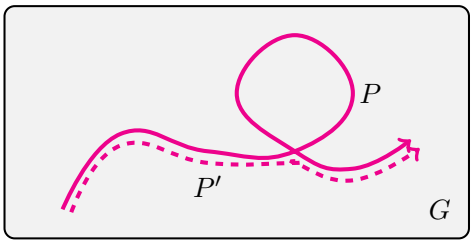
Trditev 9.5 *Naj bo G graf in u, v njegovi točki. V grafu G obstaja u-v-sprehod natanko tedaj, ko obstaja u-v-pot. Posebej, najkrajši u-v-sprehod (če kak obstaja) v grafu G je pot.*

Dokaz. Dovolj je pokazati, da je morebitni najkrajši sprehod med izbranima točkama u in v pot. Naj bo P najkrajši u-v-sprehod v grafu G. Denimo, da je $|P| = k$, in oštevilčimo točke vzdolž P:

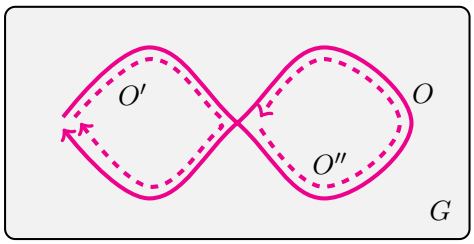
$$P = x_0 x_1 \dots x_{k-1} x_k,$$

pri čemer smo točki u in v označili z x_0 oziroma x_k. Če P ni pot, potem obstajata indeksa $i, j \in \{0, \dots, k\}$, za katera je $i < j$ in $x_i = x_j$. Stik odsekov $P_{x_0 - x_i}$ in $P_{x_j - x_k}$ je u-v-sprehod $P' = P_{x_0 - x_i} P_{x_j - x_k}$, ki je strogo krajši od P. To je nemogoče. Glej tudi sliko 9.18. □

V vsakem grafu obstaja obhod (denimo obhod dolžine 0). Ali lahko najdemo tudi cikel? Spoznajmo dva zadostna pogoja za obstoj cikla v grafu. Prvi uporabi parnost dolžine obhoda, drugi zadostno število povezav.



Slika 9.18: Če sprehod P ni pot, lahko najdemo krajši sprehod P' z istima krajiščema.



Slika 9.19: Obhod O, ki ni cikel, in krajša obhoda O' in O''.

Trditev 9.6 *Graf G vsebuje obhod lihe dolžine natanko tedaj, ko vsebuje cikel lihe dolžine. Posebej, najkrajši obhod lihe dolžine v grafu G (če kak obstaja) je cikel.*

Dokaz. Denimo, da je O najkrajši obhod lihe dolžine v grafu G. Ker obhodov dolžine 1 v grafu G ni, je $|O| \geq 3$. Z drugimi besedami, vsak obhod lihe dolžine je dolžine vsaj 3.

Točke vzdolž O oštevilčimo,

$$O = x_0 x_1 \dots x_{k-1} x_k.$$

Pri tem velja $x_0 = x_k$.

Če O ni cikel, potem obstajata različna indeksa $i, j \in \{0, \dots, k\}$, za katera je $0 < i < j$ in $x_i = x_j$. Konstruirajmo obhoda $O' = O_{x_0 - x_i} O_{x_j - x_k}$ in $O'' = O_{x_i - x_j}$. Zanju velja $|O'| \geq 1$, $|O''| \geq 1$ in $|O'| + |O''| = |O|$. Ker je O lihe dolžine, je natanko eden od O', O'' lihe dolžine, hkrati pa je tudi strogo krajši od O. To je v prostislovju s predpostavko, da je O najkrajši obhod lihe dolžine v grafu G.

Na sliki 9.19 krajša obhoda O' in O'' predstavimo črtkano. □

Trditev 9.7 *Naj bo G graf z n točkami in vsaj n povezavami. Potem G vsebuje cikel.*

Dokaz. Grafu G zaporedoma odstranjujemo točke stopenj 0 in 1. Z vsako odstranjeno točko število povezav pade za največ 1. Zato brez škode privzamemo, da so vse točke grafa G stopnje vsaj 2.

Prvo točko v_1 poti, ki jo bomo gradili induktivno (dokler ne pridelamo cikla), izberemo poljubno. Druga točka v_2 naj bo poljubna soseda točke v_1. Denimo, da smo induktivno konstruirali pot

$$v_1 v_2 \dots v_k. \tag{9.10}$$

Če je točka v_k sosednja z v_j, $j \in \{1, \dots, k-2\}$, potem je $v_j v_{j+1} \dots v_k v_j$ iskani cikel. Če pa je točka v_{k-1} edini sosed v_k vzdolž poti (9.10), potem točko v_{k+1} izberemo med preostalimi sosedami točke v_k v grafu G.

V grafu G je dolžina najdaljše poti navzgor omejena s številom točk, zato se induktivni postopek konča s ciklom. □

9.6 Povezanost grafov

Sprehodi v grafih nam omogočajo gibanje med točkami v grafu. Včasih iz izbrane točke ne moremo do želene ciljne točke, drugič lahko. Kako opisati strukturo grafa, ki hrani takšne informacije?

Relacijo *dosegljivosti* R v množici točk izbranega grafa G definiramo takole:

$$xRy \quad \text{natanko tedaj, ko v grafu } G \text{ obstaja } x\text{-}y\text{-sprehod.} \tag{9.11}$$

Po trditvi 9.5 lahko relacijo dosegljivosti enakovredno definiramo z opisom, da velja xRy natanko tedaj, ko v grafu G obstaja x-y-pot.

Trditev 9.8 *Naj bo G graf. Relacija dosegljivosti R v $V(G)$ je ekvivalenčna.*

Dokaz. Pokazati je potrebno, da je R refleksivna, simetrična in tranzitivna.

Za vsako točko v grafa G je zaporedje v z enim samim členom sprehod z začetkom in koncem v točki v. To pomeni, da je vRv. Posledično je relacija R refleksivna.

Če je uRv, potem v grafu G obstaja u-v-sprehod S. Obratni sprehod S^R ima začetek v točki v in konec v točki u, zato velja vRu. Torej je R simetrična relacija.

Denimo, da velja uRv in vRw. Torej obstajata u-v-sprehod S_1 in v-w-sprehod S_2. Njun stik S_1S_2 je u-w-sprehod, zato velja tudi uRw. S tem smo pospravili še tranzitivnost relacije R. □

Ekvivalenčna relacija dosegljivosti R razbije množico točk grafa na ekvivalenčne razrede,

$$V(G)/R = \{V_1, V_2, \ldots, V_\omega\}.$$

Točki u in v sta (vzajemno) dosegljivi natanko tedaj, ko pripadata istemu ekvivalenčnemu razredu (bloku razbitja) iz $V(G)/R$.

Induciranim podgrafom $G[V_1], G[V_2], \ldots, G[V_\omega]$ pravimo *povezane komponente* grafa G.

Nadalje, graf G je *povezan*, če ima natanko eno povezano komponento. Enakovredno, graf G je povezan, če za poljubni točki u, v v grafu G obstaja u-v-sprehod (ali celo u-v-pot, glej trditev 9.5).

V povezanem grafu lahko definiramo razdaljo[9] med točkami. *Razdalja* med točkama u in v (v grafu G), $\mathrm{dist}_G(u, v)$ (tudi samo $\mathrm{dist}(u, v)$, če je G razviden iz konteksta), je dolžina

[9] Razdaljo včasih "definiramo" tudi za nepovezane grafe. Zapis $\mathrm{dist}_G(u, v) = \infty$ pomeni, da točki u in v ne pripadata isti povezani komponenti grafa G.

najkrajšega u-v-sprehoda v grafu G.

$$\operatorname{dist}(u, v) = \min\{|P| \mid P \text{ je } u\text{-}v\text{-sprehod}\} \tag{9.12}$$

Navedimo še metrično lastnost grafovske razdalje.

Izrek 9.9 *Razdalja v povezanem grafu ustreza trikotniški neenakosti. Za poljubne tri točke u, v, w grafa G velja zveza*

$$\operatorname{dist}(u, w) \leq \operatorname{dist}(u, v) + \operatorname{dist}(v, w).$$

Dokaz. Naj bo S_1 najkrajši u-v-sprehod in S_2 najkrajši v-w-sprehod v grafu G. Njun stik S_1S_2 je u-w-sprehod dolžine $|S_1| + |S_2| = \operatorname{dist}(u, v) + \operatorname{dist}(v, w)$, saj zaradi optimalnosti velja $|S_1| = \operatorname{dist}(u, v)$ in $|S_2| = \operatorname{dist}(v, w)$.

Razdalja med u in w, $\operatorname{dist}(u, w)$, pa ni večja od dolžine sprehoda S_1S_2. □

9.7 Dvodelni grafi

Polni dvodelni grafi imajo lastnost, da lahko točke razbijemo v dva bloka, pri čemer sta točki sosedi natanko tedaj, ko pripadata različnima blokoma razbitja. Če namesto ekvivalence zahtevamo zgolj implikacijo, dobimo dvodelne grafe.

Graf G je *dvodelen*, če lahko točke grafa, množico $V(G)$, razbijemo v par blokov (A, B) tako, da krajišči vsake povezave $e = uv \in E(G)$ pripadata različnima blokoma. Velja torej $A \cup B = V(G)$, $A \cap B = \emptyset$ in

$$\forall uv \in E(G)(u \in A \wedge v \in B \text{ ali } u \in B \wedge v \in A).$$

Množicama A in B pravimo tudi *barvna razreda* dvodelnega grafa G, pogovorno pa ju bomo imenovali tudi *črne* in *bele* točke, *sode* in *lihe* točke, morda celo *leve* in *desne* točke. Dvodelni graf G tipično narišemo tako, da točke dveh barvnih razredov narišemo na dveh vzporednih premicah, vsaka povezava pa je daljica s po enim krajiščem na vsaki od premic.

Zgledi dvodelnih grafov so:

- polni dvodelni grafi $K_{m,n}$, skoraj po definiciji;
- sodi cikli C_{2k}, $k \geq 2$, barvna razreda sta kar množici *sodih* in *lihih* ostankov po modulu $2k$;
- hiperkocke Q_n, $n \geq 0$, točke-zaporedja razbijemo na *soda* (vsebujejo sodo mnogo enic) in *liha* (vsebujejo liho mnogo enic) — enostavno se je prepričati, da se števili enic v sosednjih točkah razlikujeta za natančno 1.

Vsak podgraf dvodelnega grafa G je dvodelen tudi sam, zato lahko dvodelne grafe karakteriziramo z družino prepovedanih podgrafov. Kateri pa so takšni grafi?

Za prvi zgled ni potrebno biti zelo domiseln. Polni graf K_3 gotovo ni dvodelen. Treh paroma sosednjih točk ne moremo razbiti v dva barvna razreda. Če pa polnemu grafu K_3 odstranimo katerokoli povezavo, pridelamo (polni) dvodelni graf $K_{1,2}$.

Podobno velja za vse lihe cikle. Cikel C_{2k+1}, $k \geq 1$, ni dvodelen graf — točke cikla C_{2k+1} so (lahko) ostanki pri deljenju s številom $2k+1$. Ostanke po vrsti $0, 1, \ldots$ "barvamo" s sodo in liho barvo. Če naj bo C_{2k+1} dvodelen graf, bi morali biti števili 0 in $2k$ različne parnosti, saj sta sosedi. Če pa ciklu C_{2k+1} odstranimo katerokoli izmed povezav, pridelamo pot na $2k+1$ točkah, ki je dvodelen graf.

Izkaže se, da so lihi cikli edina ovira za dvodelnost grafa.

Izrek 9.10 *Graf G je dvodelen natanko tedaj, ko ne vsebuje kakega lihega cikla kot podgrafa.*

Dokaz. Izrek je dovolj pokazati za povezane grafe. Graf G vsebuje lih cikel kot podgraf natanko tedaj, ko katera od njegovih komponent vsebuje lih cikel. Podobno je graf dvodelen natanko tedaj, ko so dvodelne vse njegove komponente.

Lihi cikli niso dvodelni grafi. Zato je dovolj pokazati, da lahko poiščemo ustrezno razbitje, če le v grafu nimamo nobenega lihega cikla.

Naj bo

$$A = \{v \in V(G) \mid \operatorname{dist}(v_0, v) \equiv 0 \pmod 2\} \quad \text{in}$$
$$B = \{v \in V(G) \mid \operatorname{dist}(v_0, v) \equiv 1 \pmod 2\}.$$

Potem je (A, B) razbitje množice $V(G)$.

Za vsako točko $x \in V(G)$ s P_x označimo (poljubno izbrano) najkrajšo x-v_0-pot. Za točke $x \in A$ je P_x sode dolžine, za točke $x \in B$ pa lihe.

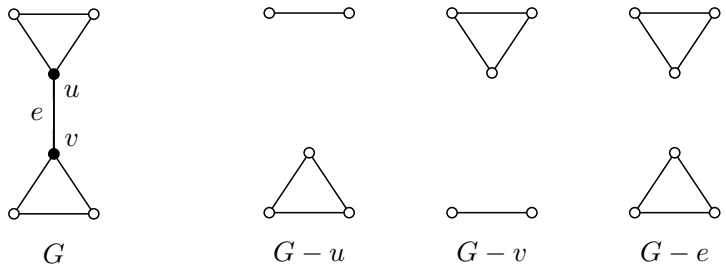
Denimo, da obstaja povezava $xy \in E(G)$, ki ima obe krajišči v istem bloku razbitja (A, B). Potem je stik sprehoda $P_y P_x^R$ s potjo-povezavo xy obhod lihe dolžine. Po trditvi 9.6 graf G vsebuje tudi cikel lihe dolžine.

Če graf G ne vsebuje nobenega lihega cikla, potem ima vsaka povezava eno krajišče v A in drugo v B. Torej sta A in B barvna razreda dvodelnega grafa G. □

Razdelek zaključimo s pojmoma prerezne povezave in prerezne točke v grafu. Povezava $e \in E(G)$ je *prerezna*, če ima graf $G-e$ *strogo več* povezanih komponent kot G. Analogno, točka $v \in V(G)$ je *prerezna*, če ima graf $G - v$ *strogo več* povezanih komponent kot G.

V povezanih grafih lahko prerezno povezavo oziroma točko opišemo na enostavnejši način. Če je G povezan, potem je $e \in E(G)$ prerezna povezava (oziroma je $v \in V(G)$ prerezna točka) natanko tedaj, ko je $G - e$ (oziroma $G - v$) nepovezan graf.

Prerezne točke in povezave so predstavljene na sliki 9.20.



Slika 9.20: Prerezni točki u, v in prerezna povezava e v grafu G.

Prerezne povezave v grafu lahko opišemo z naslednjo trditvijo.

Trditev 9.11 *Naj bo G graf. Povezava e je prerezna v grafu G natanko tedaj, ko e ne leži na nobenem ciklu grafa G.*

Dokaz. Privzeti smemo, da je G povezan graf. V nasprotnem primeru se lahko omejimo na tisto komponento grafa G, ki vsebuje povezavo e.

($\Longrightarrow$) Denimo, da je povezava $e = uv$ prerezna in vseeno leži na ciklu C_e. To po eni strani pomeni, da obstajata točki u' in v', ki nista vzajemno dosegljivi v grafu $G-e$. S P označimo u'-v'-pot v grafu G, ki mora gotovo uporabiti povezavo $e = uv$. Po drugi strani pa lahko povezavo $e = uv$ na P nadomestimo s potjo $C_e - e$ in pridelamo u'-v'-sprehod v $G - e$, kar je neumnost.

($\Longleftarrow$) Denimo, da $e = uv$ ne leži na nobenem ciklu v grafu G. Točki u in v sta v grafu G vzajemno dosegljivi, saj sta celo sosedi. Če sta dosegljivi tudi v grafu $G - e$, če torej obstaja u-v-pot P v grafu $G - e$, potem lahko v G pot P s povezavo $e = uv$ dopolnimo do cikla. To je protislovje in dokaz je zaključen. □

9.8 Drevesa in gozdovi

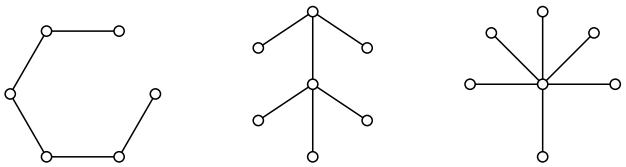
V tem razdelku obravnavamo grafe brez ciklov. Začnimo kar z ustreznima definicijama.

Drevo je povezan graf brez ciklov. *Gozd* je graf brez ciklov.

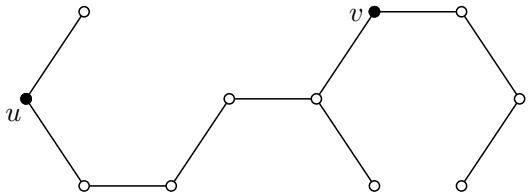
Pojma sta tesno povezana. Povezan gozd je drevo in povezane komponente gozda so drevesa. Gozd s tremi komponentami-drevesi je prikazan na sliki 9.21.

Naštejmo, zaenkrat brez dokaza, nekatere lastnosti dreves. Naj bo T poljubno drevo.

(D1) T je povezan graf.



Slika 9.21: Gozd s tremi komponentami-drevesi.



Slika 9.22: Drevo s točkama u in v.

(D2) T je brez ciklov.

(D3) $|E(T)| = |V(T)| - 1$.

(D4) Vsaka povezava $e \in E(T)$ je prerezna.

(D5) Če sta $u, v \in V(T)$, potem T vsebuje natanko eno u-v-pot.

(D6) Če sta $u, v \in V(T)$ in $uv \notin E(G)$, potem $T + uv$ vsebuje natanko en cikel.

Slika 9.22 prikazuje drevo T s točkama u in v. Drevo T ima natanko 11 točk in 10 povezav, ki so vse prerezne. V drevesu T obstaja enolično določena u-v-pot. Točki u in v nista sosedi, $T + uv$ vsebuje natanko en cikel.

Lastnosti dreves (D1), ...,(D6) utemeljimo z naslednjim izrekom.

Izrek 9.12 *Za graf G so enakovredne naslednje trditve.*

(1) *G je drevo.*

(2) *Za vsaki dve točki $u, v \in V(G)$ obstaja natanko ena u-v-pot.*

(3) *G je povezan in vsaka povezava $e \in E(G)$ je prerezna.*

(4) *G je povezan in $|E(G)| = |V(G)| - 1$.*

(5) *G ne vsebuje ciklov in $|E(G)| = |V(G)| - 1$.*

(6) *G ne vsebuje ciklov in za vsaki nesosednji točki $u, v \in V(G)$ graf $G + uv$ vsebuje natanko en cikel.*

Dokaz. (1)⇒(2) Izberimo par točk $u, v \in V(G)$, za katerega obstajata dve u-v-poti P_1 in P_2, pri čemer naj bo vsota dolžin $|P_1| + |P_2|$ najmanjša možna. V tem primeru je stik $P_1 P_2^R$ cikel.

(2)⇒(3) Graf G je povezan, saj med poljubnima točkama obstaja pot. Če obstaja povezava $e = uv$, ki ni prerezna, po trditvi 9.11 leži na ciklu. To pomeni, da obstajata vsaj dve u-v-poti v G.

(3)⇒(4) Izberimo poljubno točko v_0 in za vsako točko $v \in V(G - v_0)$ izberimo najkrajšo v-v_0-pot P_v v grafu G. Predhodnica točke v, v_p, je druga točka na poti P_v. Preslikava $\Phi : V(G - v) \to E(G)$, ki vsaki točki $v \in V(G - v)$ priredi povezavo vv_p, je *injektivna*, zato je $|E(G)| \geq |V(G)| - 1$.

Denimo, da Φ ni surjektivna, naj povezava xy ne pripada sliki preslikave Φ. To pomeni, da $x \neq y_p$ in $y \neq x_p$. V tem primeru je $P_x P_y^R$ sprehod v grafu $G - xy$ z začetkom v x in koncem v y. Odtod sledi, da povezava xy v grafu G leži na ciklu in po trditvi 9.11 ni prerezna.

(4)⇒(5) Denimo, da je G povezan, $|E(G)| = |V(G)| - 1$ in G vsebuje cikel C. Za vsako točko $v \in V(G) \setminus V(C)$ izberimo najkrajšo pot z začetkom v v in koncem v točki cikla C in jo označimo s P_v. Drugo točko na poti P_v označimo z v_p in jo imenujemo predhodnica točke v. Preslikava $\Phi : V(G) \setminus V(C) \to E(G)$, ki vsaki točki $v \in V(G) \setminus V(C)$ priredi povezavo vv_p, je *injektivna*. Število povezav grafa G je torej večje ali enako od vsote $|V(C)| + (|V(G)| - |V(C)|)$ (prvi člen označuje število povezav s C, drugi je spodnja meja za število povezav, ki niso na C), kar je protislovje.

(5)⇒(6) Izberimo graf G z n točkami, brez ciklov in z $n - 1$ povezavami. Naj bosta u, v nesosednji točki v grafu G. Denimo, da u in v pripadata različnima komponentama grafa G. Potem lahko graf G zapišemo kot disjunktno unijo grafov G_1 in G_2, pri čemer $u \in V(G_1)$ in $v \in V(G_2)$. Grafa G_1 in G_2 sta brez ciklov, zato po trditvi 9.7 velja

$$|E(G_1)| \leq |V(G_1)| - 1 \quad \text{in} \quad |E(G_2)| \leq |V(G_2)| - 1.$$

Odtod sledi

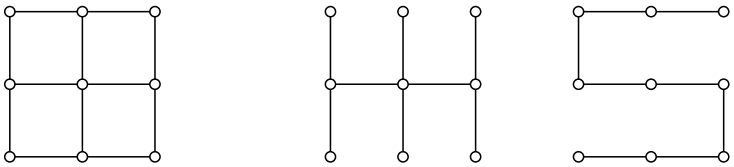
$$|E(G)| = |E(G_1)| + |E(G_2)| \leq |V(G_1)| + |V(G_2)| - 2 = |V(G)| - 2,$$

kar je nemogoče.

Torej je graf G povezan in pri poljubnih nesosednjih točkah u in v graf $G+uv$ vsebuje *vsaj* en cikel. Če vsebuje $G+uv$ dva različna cikla C_1 in C_2, potem tudi graf $(C_1 \cup C_2) - uv \subseteq G$ vsebuje cikel, kar je neumnost.

(6)⇒(1) Denimo, da G ni povezan. Potem obstajata točki u, v iz različnih komponent grafa G. Seveda u in v nista sosedi. Toda $G + uv$ vsebuje cikel skozi povezavo uv, kar pomeni, da obstaja u-v-pot v grafu G. To je protislovje, ki zaključi dokaz. □

Naj bo T drevo na vsaj dveh točkah. Ker je T povezan graf, ne vsebuje izoliranih točk, zato velja $\delta(T) \geq 1$. Je morda lahko $\delta(T) \geq 2$? Odgovor je negativen:



Slika 9.23: Graf G in dve njegovi vpeti drevesi.

Trditev 9.13 *Naj bo T drevo in $|V(T)| \geq 2$. Potem T vsebuje vsaj dva lista.*

Dokaz. Ker je T povezan graf in je $|V(T)| = n \geq 2$, je $\delta(T) \geq 1$. Denimo, da T vsebuje največ en list. Z U označimo množico $n-1$ točk drevesa T, ki so vse stopnje vsaj 2.

Po trditvi 9.1 je

$$2 \cdot |E(T)| = \sum_{v \in V(T)} \deg(v) \geq 1 + \sum_{v \in U} \deg(v) \geq 1 + 2 \cdot |U| = 2|V(G)| - 1,$$

po izreku 9.12 pa za vsako drevo velja

$$2 \cdot |E(T)| = 2|V(G)| - 2.$$

Pridelali smo protislovje, torej T vsebuje vsaj dva lista. □

Pravimo, da je T_G *vpeto drevo* v grafu G, če je T_G drevo in hkrati vpet podgraf grafa G. Zgled vpetih dreves je predstavljen na sliki 9.23.

Izrek 9.14 *Graf G vsebuje vpeto drevo natanko tedaj, ko je G povezan graf.*

Dokaz. ($\Longrightarrow$) Naj bosta u, v poljubni točki grafa G in T_G vpeto drevo. V drevesu T_G obstaja (celo enolično določena) u-v-pot P, ki je seveda prisotna tudi v večjem grafu G. Zato je G povezan.

($\Longleftarrow$) Vpeto drevo pridelamo kot zadnji graf maksimalnega (po številu členov) induktivno definiranega zaporedja grafov

$$G_0, G_1, \ldots, G_k, \tag{9.13}$$

kjer je $G_0 = G$ ter za vsak $i \in \{0, \ldots, k-1\}$ obstaja povezava $e_i \in E(G_i)$, ki v grafu G_i *ni prerezna*, in velja $G_{i+1} = G_i - e_i$.

Induktivno[10] lahko pokažemo, da so vsi grafi v zaporedju (9.13) povezani in prav tako vpeti podgrafi grafa $G = G_0$. Začetni graf G_0 je povezan po predpostavki, ravno tako pa povezanost z odstranjevanjem povezave, ki ni prerezna, ohranjamo.

[10] Algoritmično vpeto drevo raje konstruiramo z dodajanjem povezav grafa G, s katerimi ne ustvarimo nobenega cikla. Če začnemo s praznim grafom na n točkah, je potrebno dodati natanko $n-1$ "ustrezno izbranih" povezav.

Zakaj zaporedja (9.13) ne moremo podaljšati z novim členom? Edina ovira je dejstvo, da so v grafu G_k vse povezave prerezne. Po trditvi 9.11 je G_k brez ciklov in je posledično vpeto drevo v G. □

Razdelek končajmo z vprašanjem. Denimo, da je G povezan graf z n točkami. Kolikšno je največje možno število prereznih povezav oziroma prereznih točk v grafu G?

Kar se števila prereznih povezav tiče, je odgovor enostaven. Če je G drevo, potem so vse njegove povezave, ki jih je natanko $n-1$, prerezne. Če G ni drevo, jih je strogo manj.

So lahko prerezne tudi vse točke v grafu? To pa ni možno.

Trditev 9.15 *Naj bo G povezan graf in $|V(G)| \geq 2$. Potem G vsebuje vsaj dve točki, ki nista prerezni.*

Dokaz. Naj bo T_G vpeto drevo grafa G. Po trditvi 9.13 v drevesu T_G obstajata dva lista u in v (morda jih je celo več).

Po izreku 9.14 je graf $G - v$ povezan, saj je $T_G - v$ njegovo vpeto drevo. Ravno tako je povezan $G - u$. Posledično u in v nista prerezni točki grafa G. □

Vseeno sta trditvi 9.13 in 9.15 tesni. Pot P_n, $n \geq 2$, ima natančno 2 lista in $n-2$ prereznih točk.

9.9 Eulerjevi in Hamiltonovi grafi

V tem razdelku se bomo ukvarjali z iskanjem posebej dolgega sprehoda v grafu G. Če ima graf G vsaj eno povezavo uv, potem je alternirajoče zaporedje $uvuvuvuvuv$ sprehod v grafu G, ki neprestano uporablja isto povezavo in se ne loči od točk u oziroma v. S takšno izbiro lahko pridelamo poljubno dolge sprehode.

Takšni "goljufiji" se bomo izognili tako, da bomo bodisi prepovedali ponovno uporabo povezav bodisi večkratno obiskovanje točk. Osredotočili se bomo zgolj na obhode.

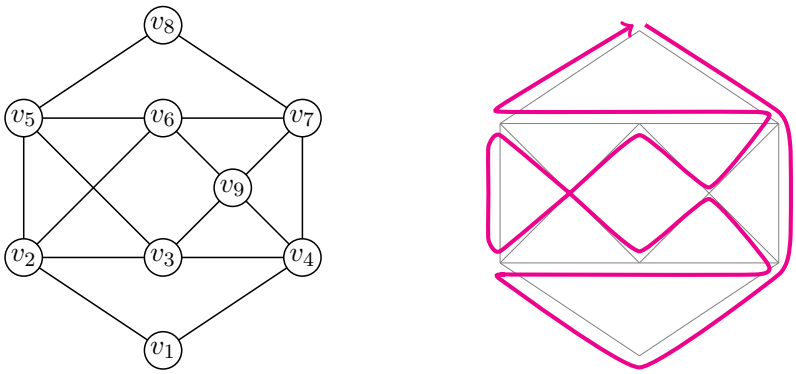
Eulerjev obhod

Spomnimo se definicije enostavnega sprehoda, tak sprehod vsako povezavo uporabi največ enkrat. Enostaven obhod v grafu, ki uporabi vse povezave, imenujemo *Eulerjev obhod*.

Dolžina vsakega enostavnega sprehoda v grafu G je navzgor omejena z $|E(G)|$. Eulerjev obhod v grafu, če obstaja, doseže to zgornjo mejo.

Graf G je *Eulerjev*, če ima kak Eulerjev obhod. Zanimalo nas bo, kako opisati Eulerjeve grafe. Oziroma, kako odločiti, ali je graf Eulerjev ali ne.

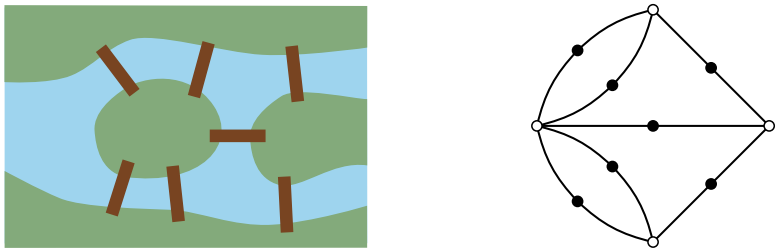
Zgled grafa in njegovega Eulerjevega obhoda predstavimo na sliki 9.24.



Slika 9.24: Eulerjev graf in njegov Eulerjev obhod $v_8v_7v_4v_1v_2v_3v_4v_9v_3v_5v_2v_6v_9v_7v_6v_5v_8$.

Odkod ime? Mesto Königsberg[11] je imelo v začetku osemnajstega stoletja sedem mostov, ki so povezovali rečna otoka in oba bregova reke Pregel. Zemljevid je predstavljen na sliki 9.25(levo). Meščani so želeli na sprehod po mestu, ki bi prečil vsakega od mostov natanko enkrat, na koncu pa bi se vrnili domov.

Zemljevid mesta predstavimo z grafom. Točke grafa ustrezajo tako delom kopnega kot mostovom. Posamezen most povežemo z obema bregovoma, na katerih stoji. Pridelamo graf s slike 9.25(desno).



Slika 9.25: Sedem mostov v Königsbergu, prvotni Eulerjev problem. Črne točke grafa predstavljajo mostove, bele točke pa dele kopnega.

Leonhard Euler je leta 1736 pokazal, da je naloga nerešljiva. Graf s slike 9.25 ne vsebuje enostavnega sklenjenega obhoda, ki vsako povezavo uporabi natanko enkrat — tak obhod bi vsakega od mostov prehodil natanko enkrat in se na koncu vrnil v začetno točko.

Še več, hkrati je izdelal tudi kriterij, potreben in zadosten pogoj, za obstoj enostavnega obhoda v grafu, ki uporabi vse povezave v grafu — Eulerjevega obhoda.

[11] Zdaj Kaliningrad, Rusija.

Izrek 9.16 (Euler) *Naj bo G graf brez izoliranih točk. Potem G vsebuje Eulerjev obhod natanko tedaj, ko je G povezan in so vse njegove točke sodih stopenj.*

Dokaz. ($\Longrightarrow$) Naj bo graf G brez izoliranih točk in naj bo

$$O = v_1 v_2 \dots v_m v_1 \tag{9.14}$$

njegov Eulerjev obhod. Izberimo poljubno točko $v \neq v_1$. Njena stopnja je soda, saj je enaka $2 \cdot |\{i \in \{1, \dots, m\} \mid v = v_i\}|$. Eulerjev obhod namreč pri vsakem prihodu v točko v uporabi natanko dve povezavi iz $E(v)$. Tudi začetna točka v_1 obhoda O je sode stopnje, saj po trditvi 9.3 graf G ne more imeti samo ene točke lihe stopnje.

Izberimo poljubni točki u, v. Ker u in v nista izolirani točki, obhod O vsebuje obe. Odsek obhoda O med u in v je u-v-sprehod, zato je G povezan..

($\Longleftarrow$) Težja smer dokaza sledi zelo enostavni ideji. Privzemimo, da G vsebuje vsaj eno povezavo (sicer ni česa dokazovati) in z

$$O = v_1 v_2 \dots v_\ell \tag{9.15}$$

označimo najdaljši enostaven sprehod v grafu G. Pokazali bomo, da je O Eulerjev obhod. Pa začnimo.

(1) O je obhod, $v_1 = v_\ell$.

Če O ni obhod, potem njegovi krajišči ne sovpadata. Če je v konec O, potem O uporabi liho mnogo povezav iz $E(v)$. Ker je $|E(v)| = \deg(v)$ sodo število, obstaja povezava $e \in E(v)$, ki je O ne uporabi. Torej lahko sprehod O nadaljujemo vzdolž e, kar nas pripelje v protislovje s predpostavko, da je O maksimalne dolžine.

(2) Če O vsebuje točko v, potem uporabi vse povezave iz $E(v)$.

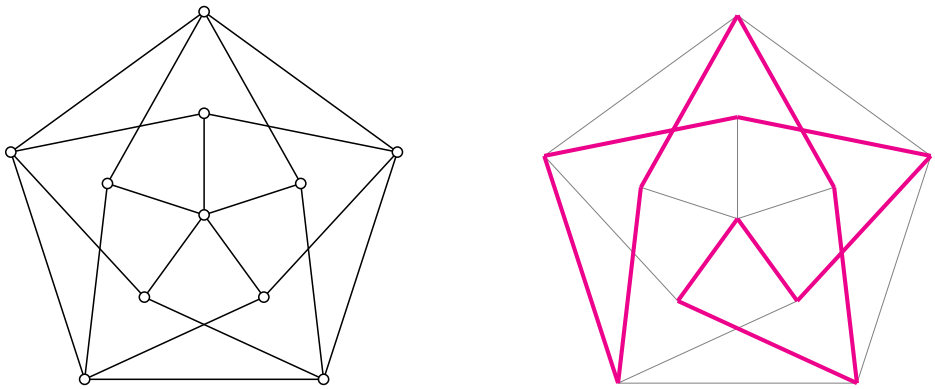
Denimo, da obhod (uporabimo (1)) O vsebuje točko v in ne uporabi vseh povezav iz $E(v)$. To pomeni, da obstaja indeks $i \in \{1, \dots, \ell\}$, za katerega je $v = v_i$. Obhod $O' = O_{v_i - v_\ell} O_{v_1 - v_i}$ dobimo s premikom začetne točke, je iste dolžine kot O in se začne in konča v točki $v = v_i$. Če O ne uporabi vseh povezav iz $E(v)$, potem lahko obhod O' podaljšamo vzdolž ene od neuporabljenih povezav iz $E(v)$ in dobimo daljši enostaven sprehod. To je v nasprotju s predpostavkami.

(3) O vsebuje vse točke grafa G.

Povezave vzdolž O "obarvajmo" z rdečo barvo, preostale pa s črno. Po (2) za vsako točko v iz obhoda O velja, da so vse povezave iz $E(v)$ rdeče. Če pa O ne gre skozi točko x, potem so vse povezave iz $E(x)$ črne.

Če O ne vsebuje vseh točk grafa G, potem obstajata tako točka v, ki se dotika samih rdečih povezav, in točka x, ki se dotika samih črnih. Ker je G povezan, obstaja tudi v-x-pot P.

Začetna povezava vzdolž P je rdeča in končna povezava je črna. To pomeni, da obstaja točka y na poti P, ki se dotika tako rdeče kot črne povezave. To pa je v nasprotju z (2).



Slika 9.26: Grötzschev graf G in Hamiltonov cikel v G.

Pod črto: Najdaljši enostaven sprehod O v povezanem grafu G (s samimi sodimi stopnjami točk in brez izoliranih točk) je obhod (1), vsebuje vse točke (3) in z vsako točko uporabi tudi vse povezave, ki se v tej točki stikajo (2). Torej je O Eulerjev obhod. □

Karakterizacija Eulerjevih grafov se je izkazala za relativno enostaven problem. Vsakega od možnih odgovorov na vprašanje "*Ali je graf G Eulerjev?*" lahko učinkovito utemeljimo: Pozitivni odgovor kar s konstrukcijo kakega Eulerjevega obhoda, negativnega pa tako, da v grafu G poiščemo točko lihe stopnje ali pa pokažemo, da ima G več netrivialnih komponent.

Navidezno sorodni problem Hamiltonovega cikla, ki ga bomo spoznali v preostanku razdelka, se bo izkazal za bistveno težjega.

Hamiltonov cikel

Hamiltonov cikel v grafu G je cikel, ki vsebuje vse točke grafa G. Z drugimi besedami, Hamiltonov cikel je vpet cikel v grafu G.

Graf G je *Hamiltonov*, če vsebuje kak Hamiltonov cikel. Pri obravnavi Hamiltonovih grafov se brez težav omejimo na povezane grafe z vsaj tremi točkami.

Grötszchev graf, predstavljen kot zgled na sliki 9.26, je Hamiltonov.

Odkod ime? William Rowan Hamilton je bil irski matematik (in fizik) iz devetnajstega stoletja. Raziskovalno se s teorijo grafov ni ukvarjal, je pa izumil namizno igro, v kateri je cilj poiskati cikel skozi vse točke grafa — Hamiltonov cikel. Sama igra komercialno ni bila uspešna, ime pa se je vseeno prijelo.

Karakterizacija Hamiltonovih grafov se zdi[12] mnogo težavnejša kot karakterizacija Eu-

[12] "P *je* NP *ali* P *ni* NP?" je verjetno najpomembnejše odprto vprašanje v teoretičnem računalništvu.

lerjevih grafov. Vprašanje "*Ali je graf G Hamiltonov?*" ima seveda dva možna odgovora. Pozitivni odgovor relativno enostavno utemeljimo s konstrukcijo kakega Hamiltonovega cikla. Negativnega pa v splošnem primeru na *enostaven*[13] način ne bomo znali utemeljiti.

Vseeno bomo spoznali en potreben pogoj kot tudi en zadosten pogoj za obstoj Hamiltonovega cikla v grafu. Vendar pa pogoja v vmesnem prostoru puščata bogat razred grafov, ki izpolnjujejo potrebni pogoj (in bi lahko imeli Hamiltonov cikel), ne izpolnjujejo pa zadostnega pogoja (in Hamiltonovega cikla ne vsebujejo nujno). V zaključku razdelka se bomo spopadli s Petersenovim grafom, grafom iz vmesnega prostora.

Potrebni pogoj

Ali lahko Hamiltonov graf G vsebuje prerezno točko?

Denimo, da je v prerezna točka grafa G in obhod C vsebuje vse točke grafa. Vsak premik obhoda C iz ene komponente grafa $G - v$ v drugo gre skozi točko v. Ker sta komponenti grafa $G - v$ vsaj dve, mora C točko v uporabiti vsaj dvakrat. Torej C ni cikel.

Grafi s prereznimi točkami ne vsebujejo Hamiltonovih ciklov.

Slika 9.27 prikazuje graf G s točkama u, v, pri čemer ima graf $G - u - v$ tri povezane komponente, označimo jih z G_1, G_2, G_3.

Če obhod C vsebuje vse točke grafa G, potem opravi *vsaj* tri prehode med G_1, G_2 in G_3. Pri takšnem premiku obišče tudi eno od točk u, v. Obhod C uporabi točki u in v skupaj vsaj trikrat, vsaj eno torej vsaj dvakrat. C torej ni cikel.

Tudi graf G s slike 9.27 nima Hamiltonovega cikla, sklep smo izpeljali podobno kot v primeru prerezne točke.

To (še toplo) idejo lahko posplošimo tudi na večje podmnožice točk.

Izrek 9.17 (potrebni pogoj za obstoj Hamiltonovega cikla) *Naj bo G Hamiltonov graf. Potem je G povezan in za vsako neprazno množico točk $U \subseteq V(G)$ moči k velja, da $G - U$ vsebuje največ k povezanih komponent.*

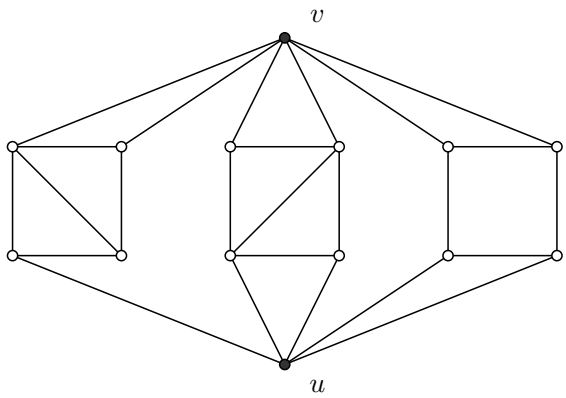
Dokaz. Naj bo C Hamiltonov cikel v grafu G. Odseki vzdolž C so poti med vsemi možnimi pari točk iz G, zato je G povezan.

Izberimo neprazno podmnožico U točk grafa G. Število povezanih komponent grafa $G - U$ je manjše ali enako od števila povezanih komponent grafa $C - U$, saj je $C - U$ vpet podgraf grafa $G - U$.

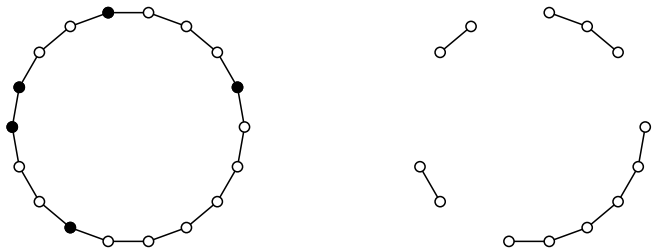
Število povezanih komponent grafa $C - U$ pa je manjše ali enako $|U|$. To velja pri

V primeru, ko razreda računske zahtevnosti P in NP nista enaka (empirične izkušnje kažejo v to smer), ne moremo računati na enostaven način dokazovanja, da posamezen graf nima Hamiltonovega cikla.

[13] Preverimo lahko vsa možna zaporedja točk grafa. Toda število takšnih zaporedij z naraščajočim številom točk raste prehitro, da bi smeli tak način imeti za enostavnega.



Slika 9.27: Graf G brez Hamiltonovega cikla, $G - u - v$ ima tri komponente..



Slika 9.28: Če iz cikla odstranimo 5 točk, razpade na največ pet komponent (lahko celo manj).

$|U| = 1$, če ciklu odstranimo eno točko, pridelamo pot. Z vsako dodatno odstranjeno točko pa število povezanih komponent naraste za največ 1.

Glej tudi sliko 9.28. □

Potrebni pogoj za obstoj Hamiltonovega cikla, izrek 9.17, ponavadi beremo v kontrapoziciji. Če graf G vsebuje "majhno" množico točk U, za katero ima $G - U$ "preveč" povezanih komponent (vsaj $|U| + 1$), potem G nima Hamiltonovega cikla.

Izrek 9.17 ima v družini dvodelnih grafov enostavno posledico.

Trditev 9.18 *Naj bo G dvodelni graf in A, B njegova barvna razreda. Če je $|A| \neq |B|$, potem G ni Hamiltonov.*

Dokaz. Privzamemo lahko, da je $|A| < |B|$. Graf $G - A$ je prazen in ima $|B|$ komponent (izoliranih točk). □

Zadostni pogoj

Opazujmo graf z $n \geq 3$ točkami. Hamiltonov cikel bomo gotovo lažje poiskali z naraščajočim številom povezav, kot tudi z večanjem stopenj točk.

Poskusimo poiskati graf s kar se da veliko *minimalno* stopnjo točk v grafu (veliko v primerjavi s številom točk grafa), ki nima Hamiltonovega cikla. Graf G_k konstruirajmo takole: disjunktni uniji dveh kopij polnega grafa K_k dodajmo univerzalno[14] točko u. Graf G_k ima natanko $2k+1$ točk in $\delta(G_k) = k = \left\lfloor \frac{|V(G_k)|}{2} \right\rfloor$. Ker je u prerezna točka, graf G_k ni Hamiltonov.

V nadaljevanju bomo pokazali, da je že malenkost večja minimalna stopnja (*navzgor* zaokrožena polovica števila točk) dovolj za obstoj Hamiltonovega cikla.

Začnimo s tehničnim rezultatom.

Izrek 9.19 (Bondy-Chvátal) *Naj bo G graf z $n = |V(G)| \geq 3$ točkami in u, v nesosednji točki, za kateri velja $\deg_G(u) + \deg_G(v) \geq n$. Potem je G Hamiltonov natanko tedaj, ko je $G + uv$ Hamiltonov.*

Dokaz. Z dodajanjem povezave uv lastnosti hamiltonskosti grafa gotovo ne moremo izgubiti. Privzemimo preostalo-slabo možnost: $G+uv$ je Hamiltonov in G ni Hamiltonov graf.

Hamiltonov cikel C v grafu $G + uv$ v tem primeru gotovo uporabi povezavo uv. Točke grafa G oštevilčimo vzdolž Hamiltonovega cikla $C = v_1 v_2 \dots v_n$, pri čemer naj velja $v_1 = u$ in $v_n = v$.

Množico indeksov sosed točke $u = v_1$ označimo z I,

$$I = \{i \mid v_i \sim_G u\}.$$

Z J pa označimo množico naslednikov indeksov sosed točke $v = v_n$,

$$J = \{j + 1 \mid v_j \sim_G v\}.$$

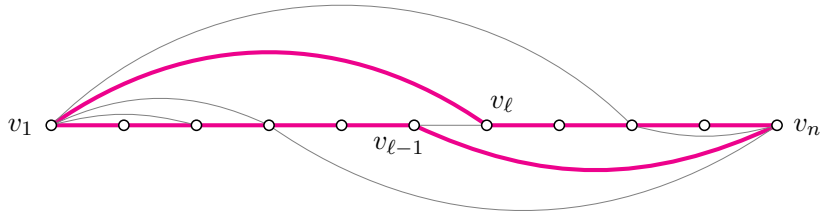
Tako I kot J sta vsebovani v $\{2, \dots, n\}$, zato je $|I \cup J| \leq n - 1$. Po drugi strani je $|I| + |J| = \deg_G(u) + \deg_G(v) \geq n$. Zato je

$$|I \cap J| = |I| + |J| - |I \cup J| \geq n - (n-1) \geq 1.$$

Torej obstaja indeks $\ell \in I \cap J$. Točka v_ℓ je soseda točke v_1 in $v_{\ell-1}$ je soseda točke v_n.

Hamiltonov cikel v G lahko konstruiramo z uporabo povezav $v_1 v_\ell$, $v_{\ell-1} v_n$ in odsekov Hamiltonovega cikla C. Glej sliko 9.29. Pridelamo protislovje, ki zaključuje dokaz. □

[14] Točka v grafu je univerzalna, če je soseda z vsemi preostalimi točkami grafa.



Slika 9.29: Hamiltonov cikel v grafu G.

Diracov izrek, enostavna posledica izreka 9.19, določi tesno spodnjo mejo minimalne stopnje, s katero zagotovimo Hamiltonov cikel.

Izrek 9.20 (Dirac) *Naj bo G graf z $n = |V(G)| \geq 3$ točkami. Če je $\delta(G) \geq \frac{n}{2}$, potem je G Hamiltonov.*

Dokaz. Po izreku 9.19 lahko grafu G korakoma dodamo vse povezave med pari nesosednjih točk, pri tem pa lastnost (ne)hamiltonskosti ohranjamo. V končni fazi pridelamo poln graf K_n, ki je Hamiltonov. Posledično je Hamiltonov tudi prvotni graf G. □

Petersenov graf ni Hamiltonov

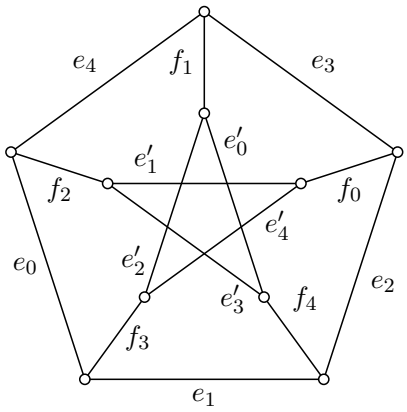
Razdelek končajmo z obravnavo Petersenovega grafa, grafa, ki je v najrazličnejših grafovskih problemih velikokrat pod drobnogledom. Petersenov graf P, predstavljen na sliki 9.30, je kubičen graf na 10 točkah, ki premore bogato simetrijo. Na sami risbi je razvidna petkotniška simetrija — če graf P zavrtimo za petino polnega kota ali pa prezrcalimo preko navpične premice, pridelamo avtomorfizem[15] Petersenovega grafa P.

Petersenov graf ne zadošča predpostavkam (Diracovega) izreka 9.20. Po drugi strani pa, brez dokaza, z odstranjevanjem majhne množice točk ne razpade na preveč komponent.

Pokazali bomo, da Petersenov graf P ni Hamiltonov. Metoda bo posebej prirejena Petersenovemu grafu in je za druge grafe ne bomo mogli uporabiti.

Povezave grafa P, v skladu z oznakami na sliki 9.30, označimo za *zunanje* (to so povezave $e_0, \ldots, e_4$), *notranje* ($e'_0, \ldots, e'_4$) in *vmesne* ($f_0, \ldots, f_4$).

[15] Avtomorfizem grafa je izomorfizem grafa samega vase — permutacija točk grafa, ki ohranja sosednost.

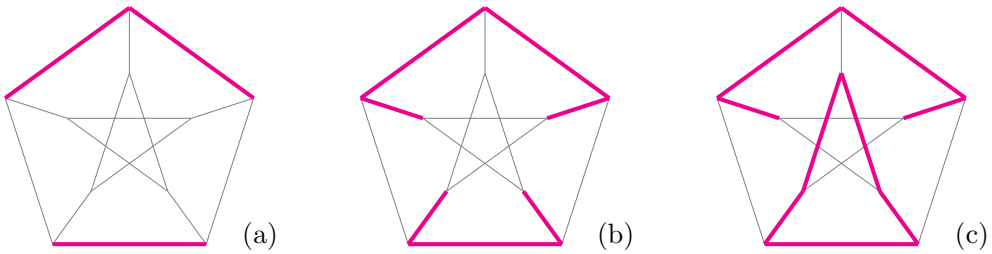


Slika 9.30: Petersenov graf P z označenimi povezavami.

Denimo, da je C_P Hamiltonov cikel v P. Cikel C_P vsebuje nekaj zunanjih povezav, ki se morajo dotikati vseh zunanjih točk. Torej uporabi vsaj tri zunanje povezave. Še več, če uporabi natanko tri izmed zunanjih povezav, te ne smejo ležati zaporedno.

Pri tem nikakor ne more porabiti vseh petih (saj C_P ne vsebuje petcikla), ravno tako pa tudi ne natanko štirih. V tem primeru bi C_P uporabil natanko dve zaporedni vmesni povezavi — in vseh pet notranjih povezav, ki se dotikajo neuporabljenih vmesnih povezav. To bi skupaj zneslo 11 povezav.

Brez škode za splošnost lahko torej privzamemo, da C_P med zunanjimi povezavami uporabi natanko e_1, e_3 in e_4, glej sliko 9.31(a). Posledično uporabi tudi vmesne povezave f_0, f_2, f_3 in f_4 (slika 9.31(b)) in zato tudi notranji povezavi e'_0 in e'_2, glej sliko 9.31(c). Posledično C_P vsebuje cikel dolžine 5, kar ni mogoče.



Slika 9.31: Petersenov graf P ne vsebuje Hamiltonovega cikla.

9.10 Barvanje točk

V zadnjem razdelku predstavimo problem barvanja točk grafov. Barvanje točk grafa je preslikava iz množice točk grafa v množico barv, pri kateri sosednjih točk ne obarvamo z isto barvo[16].

Množico barv lahko nadomestimo z množico naravnih števil, ravno tako lahko omejimo število barv na zalogi. Preslikava

$$c : V(G) \to \{1, \ldots, k\} \tag{9.16}$$

je *k-barvanje* točk grafa G, če

(COL) za vsako povezavo $uv \in E(G)$ velja $c(u) \neq c(v)$.

Graf G je *k-obarvljiv*, če ima kako k-barvanje. Najmanjše naravno število k, za katerega je graf k-obarvljiv, je *kromatično število* grafa G in ga označimo s $\chi(G)$.

Pri barvanju točk se smemo omejiti na obravnavo povezanih grafov. Povezav, ki predstavljajo omejitve za izbiro barv, med različnimi komponentami grafa ni. To pomeni, da lahko barve točk v posamezni komponenti izbiramo neodvisno od preostalih komponent.

Mnoge kombinatorične in optimizacijske probleme je moč prepisati kot problem k-barvanja grafov ali celo iskanja kromatičnega števila grafa. Naštejmo nekaj primerov.

Oddajnikom bi radi dodelili frekvence, pri čemer bližnjim parom oddajnikom zaradi možne interference signalov ne moremo dodeliti istih frekvenc. Pri tem bi radi, pasovna širina je draga, uporabili kar se da malo frekvenčnih pasov.

Sestaviti želimo urnik predavanj na fakulteti. Trojicam *predavatelj-predmet-skupina študentov* prirejamo barve (pare *ura-dan*). Pri tem pa niti istega predavatelja niti iste skupine študentov ne moremo hkrati napotiti v več predavalnic. Urnik bi radi naredili kar se da kompakten. Majhno število porabljenih barv pomeni, da predavanja ne bodo potekala od jutra do večera. Maksimalno število točk iste barve pa ustreza številu potrebnih prostorov.

Tudi igro *Sudoku* si lahko predstavljamo kot problem 9-barvanja grafa, v katerem so nekatere točke vnaprej predpisanih barv. V kvadratke igralne plošče velikosti 9×9 želimo razporediti števke $1, \ldots, 9$, pri čemer v nekatere pare kvadratov (lahko bi jim rekli sosednji) ne smemo hkrati zapisati iste števke.

Poiščimo kromatična števila za nekatere osnovne družine grafov.

Naj bo G graf z n točkami, s c označimo barvanje točk grafa G.

(1) $\chi(G) = 1$ natanko tedaj, ko je G brez povezav.

[16] V literaturi se pojavljata dva pristopa. Nekateri avtorji kot barvanje točk grafa definirajo poljubno preslikavo iz množice točk v množico barv in imenujejo barvanje *pravilno*, če imata sosednji točki vselej različni barvi. Sami bomo, pri tem nismo edini, že v samo definicijo barvanja vtkali zahtevo o različnih barvah sosednjih točk. Z *nepravimi* barvanji, ki dopuščajo isto barvo tudi za sosednji točki, se niti ne bomo ukvarjali.

Če je graf G brez povezav, lahko vse točke grafa G obarvamo z isto barvo. Če pa je $uv \in E(G)$, potem je $c(u) \neq c(v)$, zato ena sama barva ni dovolj.

(2) $\chi(G) \leq 2$ natanko tedaj, ko je G dvodelen.

Naj bosta A in B barvna razreda dvodelnega grafa G. Vse točke iz A obarvamo z barvo 1, tiste iz B pa z 2. Graf G je 2-obarvljiv. Obratno, množici točk barve 1 oziroma barve 2 sta barvna razreda grafa G. Vsaka povezava ima namreč krajišči različnih barv.

(3) $\chi(G) \leq n$; natančneje, $\chi(G) = n$ natanko tedaj, ko je G poln graf; $\chi(K_n) = n$.

Če je c bijekcija iz $V(G)$ v $\{1, \ldots, n\}$, potem je c tudi barvanje točk. V primeru, ko G vsebuje par nesosednjih točk u, v, lahko točki u in v obarvamo z isto barvo. Če vse preostale točke obarvamo s samimi različnimi barvami, dobimo $(n-1)$-barvanje točk grafa G.

(4) Če je T drevo z vsaj eno povezavo, potem je $\chi(T) = 2$.

Drevo je graf brez ciklov. Zato je T dvodelen graf.

(5) $\chi(Q_n) = 2$, če je $n \geq 1$.

Tudi hiperkocka Q_n je dvodelen graf.

(6) $\chi(C_{2k}) = 2$ in $\chi(C_{2k-1}) = 3$, za vse $k \geq 2$.

Sodi cikli so dvodelni grafi (in imajo vsaj eno povezavo), zato je $\chi(C_{2k}) = 2$. Lihi cikli niso dvodelni, torej je $\chi(C_{2k-1}) \geq 3$. So pa 3-obarvljivi, saj lahko izbrano točko lihega cikla obarvamo z barvo 3, preostale, ki inducirajo pot, pa alternirajoče z barvama $1, 2$.

Meje za kromatično število in požrešno barvanje točk

Kako se lahko enostavno prepričamo, da je kromatično število grafa veliko? Z $\omega(G)$ označimo velikost največjega polnega podgrafa v grafu G. Če največji polni podgraf v G vsebuje k točk, potem že zanje potrebujemo vsaj k različnih barv.

Trditev 9.21 *Za vsak graf G je $\chi(G) \geq \omega(G)$.*

Zgornjo mejo za kromatično število grafa upravičimo s konstrukcijo barvanja točk.

Točke grafa G uredimo v zaporedje

$$\xi_G = (x_1, x_2, \ldots, x_n). \tag{9.17}$$

Požrešno barvanje točk grafa G glede na vrstni red točk ξ_G je barvanje $c : V(G) \to \mathbb{N}$, ki ga konstruiramo induktivno:

> Privzemimo, da smo obarvali točke $v_1, v_2, \ldots, v_{i-1}$.
>
> Barva točke v_i, $c(v_i)$, je *najmanjše* pozitivno naravno število, ki *ni uporabljeno* kot barva že-obarvanih-sosed točke v_i. (9.18)

Za vsak $i \in \{1, \ldots, n\}$ velja:

$$c(v_i) = \min\{x \in \mathbb{N}^+ \mid x \notin \{c(v_1), \ldots, c(v_{i-1})\}\} \tag{9.19}$$

Izrek 9.22 *Požrešno barvanje, ne glede na vrstni red točk grafa, vedno uporabi največ* $\Delta(G)+1$ *barv. Za vsak graf* G *je* $\chi(G) \leq \Delta(G)+1$.

Dokaz. Z n označimo število točk grafa G in izberimo vrstni red točk ξ. Naj bo c požrešno barvanje glede na ξ.

Ocenimo število uporabljenih barv sosed točke v_i v trenutku, ko je v_i na vrsti za barvanje:

$$|\{c(v_j) \mid j < i \text{ in } v_j \sim v_i\}| \leq |\{v_j \mid j < i \text{ in } v_j \sim v_i\}| \leq \deg(v_i) \leq \Delta(G) \qquad (9.20)$$

Zato je vsaj ena izmed barv iz $\{1, \ldots, \Delta(G)+1\}$ v trenutku izbire barve točke v_i prosta (neuporabljena na obarvanih sosedah) za točko v_i.

Premislek velja za vsako točko zaporedja ξ, zato je požrešno barvanje $(\Delta(G)+1)$-barvanje. □

Ali pri požrešnem barvanju (9.18) res ne moremo brez barve številka $\Delta(G)+1$? Natančen pogled na oceno (9.20) pove, da barvo $\Delta(G)+1$ potrebujemo samo v primeru, ko barvamo točko v_i, ki je (i) stopnje $\Delta(G)$, (ii) ima vse sosede že obarvane in (iii) so njene sosede samih različnih barv.

Število uporabljenih barv pri požrešnem barvanju grafa G je (v splošnem) močno odvisno od vrstnega reda točk ξ. Če je vrstni red točk nesrečno izbran, lahko pri požrešnem barvanju grafa G uporabimo precej več kot $\chi(G)$ barv[17].

Če, denimo, požrešno barvamo točke cikla C_6 tako, da najprej pobarvamo par točk v_0, v_3, ki sta na medsebojni razdalji 3, bomo uporabili 3 barve. Kromatično število cikla pa je manjše, $\chi(C_6) = 2$.

Za cikle lihe dolžine C_{2k+1} velja zveza $\chi(C_{2k+1}) = 3 = \Delta(C_{2k+1}) + 1$. Tudi za polne grafe je $\chi(K_n) = n = \Delta(K_n) + 1$. Brooksov izrek poskrbi za obrat, kromatično število povezanega grafa je večje od maksimalne stopnje samo v primerih lihih ciklov in polnih grafov.

Izrek 9.23 (Brooks) *Naj bo* G *povezan graf. Če* G *ni cikel lihe dolžine niti poln graf, potem je* $\chi(G) \leq \Delta(G)$.

Dokaz. Naj bo G povezan graf, ki ni niti lih cikel niti poln graf. Velja torej $\Delta(G) \geq 2$.

Če je $\Delta(G) = 2$, potem je G bodisi sod cikel bodisi pot na vsaj treh točkah. Sodi cikli in poti so dvodelni grafi, njihovo kromatično število je enako 2.

V nadaljevanju privzamemo, da je $\Delta(G) \geq 3$. Barvanje bomo zgradili požrešno z izbiro ustreznega vrstnega reda točk.

[17] Res pa je tudi, da lahko $\chi(G)$-barvanje grafa konstruiramo požrešno, če smo pripravljeni preskusiti *vsa* zaporedja točk grafa G.

Če v povezanem grafu izberemo zaporedje točk ξ_x, v katerem točke uredimo po padajoči razdalji do x, potem ima vsaka točka $v \in V(G) \setminus \{x\}$ v trenutku požrešnega določanja njene barve vsaj enega neobarvanega soseda (vsaj eden od njenih sosedov je strogo bliže točki x in leži kasneje v zaporedju ξ_x). To pomeni, da vse točke, razen morda zadnje točke x, obarvamo z eno od barv $\{1, \ldots, \Delta(G)\}$.

Zdaj pa ločimo tri primere:

(1) G ima točko v, za katero je $\deg(v) < \Delta(G)$.

Požrešno barvanje glede na ξ_v tudi za zadnjo točko v uporabi eno od barv iz $\{1, \ldots, \Delta(G)\}$, saj ima v strogo manj kot $\Delta(G)$ sosed.

(2) G ima prerezno točko.

Naj bo v prerezna točka in $G_1, G_2, \ldots$ komponente grafa $G - v$. Z G_i^+ označimo podgraf grafa G, induciran na točkah $V(G_i) \cup \{v\}$. Točka v je v grafu G_i^+ stopnje strogo manj kot $\Delta(G)$, zato ima graf G_i^+ barvanje c_i z $\Delta(G)$ barvami (določimo ga lahko požrešno v skladu z (1)). S permutacijami barv lahko dosežemo, da barvanja $c_1, c_2, \ldots$ točko v obarvajo z isto barvo. Njihova unija je barvanje točk grafa G z $\Delta(G)$ barvami.

(3) G je regularen graf brez prerezne točke.

Vse točke grafa G imajo torej stopnjo enako $\Delta(G) \geq 3$. Ker G ni poln graf, lahko najdemo trojico točk

$$y, x, z, \tag{9.21}$$

za katero velja

(BR1) $xy, xz \in E(G)$ in $yz \notin E(G)$ ter je

(BR2) $G - y - z$ povezan graf.

Točke, ki zadoščajo (BR1) so denimo tri zaporedne točke na najkrajši poti med dvema nesosednjima točkama.

Za (BR2) je potrebno nekaj več truda.

Naj bo y', x', z' trojica točk, ki zadošča (BR1). Če je graf $G - y' - z'$ nepovezan, potem v grafu G obstajajo pari točk u, v, za katere je $G - u - v$ nepovezan. Tak je denimo par y', z'.

Izberimo par točk $x, \bar{x}$ (ki nista nujno na razdalji 2), za katerega ima najmanjša komponenta grafa $G - x - \bar{x}$ najmanjše možno število točk. Najmanjšo komponento grafa $G - x - \bar{x}$ označimo z G_0, preostale pa z G_i, $i \geq 1$.

Ker nobena od točk $x, \bar{x}$ ni prerezna v grafu G, ima vsaka od $x, \bar{x}$ soseda v vsaki od komponent grafa $G - x - \bar{x}$. Ker je graf G regularen s stopnjo vsaj 3, vsebuje G_0 vsaj dve točki. Še več, zaradi minimalnosti G_0 imata tako x kot $\bar{x}$ celo dva soseda v G_0 (če je, denimo, $\bar{\bar{x}}$ edini sosed točke $\bar{x}$ v G_0, potem ima $G - x - \bar{\bar{x}}$ komponento z manj kot $|V(G_0)|$ točkami).

Naj bo y poljubna soseda točke x iz G_0 in z soseda točke x iz G_1. Trdimo, da trojica y, x, z zadošča (BR1) in (BR2). Točki y in z nista sosedi, saj pripadata različnima komponentama grafa $G - x - \bar{x}$, imata pa skupno sosedo x. Pogoj (BR1) je izpolnjen.

Za vsako točko $v \in V(G - y - z)$ obstaja pot od v do vsaj ene od točk x ali $\bar{x}$ (v nasprotnem primeru bi bila y ali z prerezna točka v G). Po drugi strani obstaja tudi x-$\bar{x}$-pot v $G-y-z$, ki ima vse notranje točke iz G_0 (sicer je y prerezna točka *komponente* G_0 in ima $G-x-y$ komponento z manj kot $|V(G_0)|$ točkami). Torej je $G-y-z$ povezan graf. Izpolnjen je tudi pogoj (BR2).

Naj bo torej y, x, z trojica točk, ki zadošča (BR1) in (BR2). Zaporedje točk $\xi_{y,x,z}$ v grafu G konstruirajmo tako, da na prvi dve mesti postavimo točki y in z, nadaljujemo pa s preostalimi točkami, urejenimi padajoče glede na razdaljo do točke x v povezanem grafu $G - y - z$.

Požrešno barvanje c nesosednji točki y in z obarva z barvo 1, preostale točke, ki niso enake x, pa imajo vse vsaj enega soseda, ki leži kasneje v zaporedju $\xi_{y,x,z}$. Obarvamo jih z eno izmed barv iz $\{1, \ldots, \Delta(G)\}$. Zadnja točka x pa ima dve sosedi, y in z, ki sta obarvani z isto barvo. Torej barvanje c na sosedah točke x uporabi kvečjemu $\Delta(G) - 1$ različnih barv. Vsaj ena izmed $\{1, \ldots, \Delta(G)\}$ je dopustna izbira za $c(x)$. □

Barvanje Petersenovega in Grötzchevega grafa

Za sam konec bomo poiskali kromatično število (in tudi barvanje) Petersenovega in Grötzschevega grafa, P in G, ki smo ju spoznali v prejšnjem razdelku. Predstavljena sta na slikah 9.30 in 9.26. Opazimo, da imata tako P kot G petkotniško simetrijo, vrtenje točk za petino polnega kota kot tudi zrcaljenje preko navpičnice sta avtomorfizma obeh grafov.

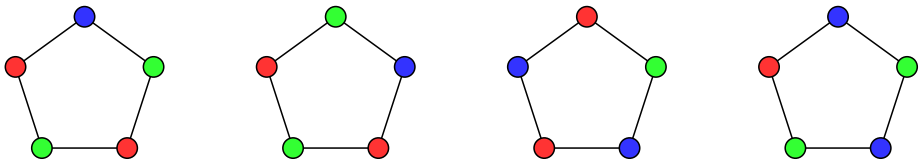
Začnimo s 3-barvanjem c cikla C_5. Ker je $\chi(C_5) = 3$, barvanje c uporabi vse tri barve. Ker je $|V(C_5)| < 6$, mora c eno izmed barv uporabiti na eni sami točki, imenujmo jo v. Preostali barvi izmenjaje uporabi na poti $C_5 - v$.

Cikel C_5 narišemo kot pravilni petkotnik in z rotacijo poskrbimo, da je točka v na vrhu. Njeno (ekskluzivno) barvo poimenujmo *modra*[18], preostali dve barvi pa *rdeča* in *zelena*. Z morebitnim zrcaljenjem preko navpičnice lahko dosežemo, da točka zelene barve sledi točki v v smeri urinega kazalca na sliki grafa C_5.

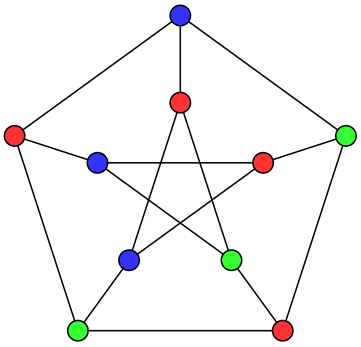
Na sliki 9.32 so predstavljena različna barvanja cikla C_5, ki jih lahko s simetrijami petkotnika in morebitnim preimenovanjem barv pretvorimo v skrajno levega — imenujmo ga *odlikovano* barvanje.

Petersenov graf P obdelamo relativno enostavno. Ker je $\omega(P) = 2$ in $\Delta(P) = 3$, po trditvi 9.21 in (Brooksovem) izreku 9.23 velja $2 \leq \chi(P) \leq 3$. Ker P ni dvodelen graf

[18] Za barve točk smo izbirali naravna števila. Preklop v barvanje z "barvami" miselno ni preveč težaven.



Slika 9.32: Različna barvanja cikla C_5 s tremi barvami, odlikovano barvanje je skrajno levo.



Slika 9.33: Petersenov graf P ima kromatično število enako 3.

(saj očitno vsebuje petcikel), je $\chi(P) \geq 3$.

Torej je $\chi(P) = 3$. Ustrezno 3-barvanje c_P sestavimo relativno brez težav, glej sliko 9.33, saj lahko za "zunanjih" pet točk Petersenovega grafa brez škode za splošnost uporabimo odlikovano barvanje petcikla s slike 9.32.

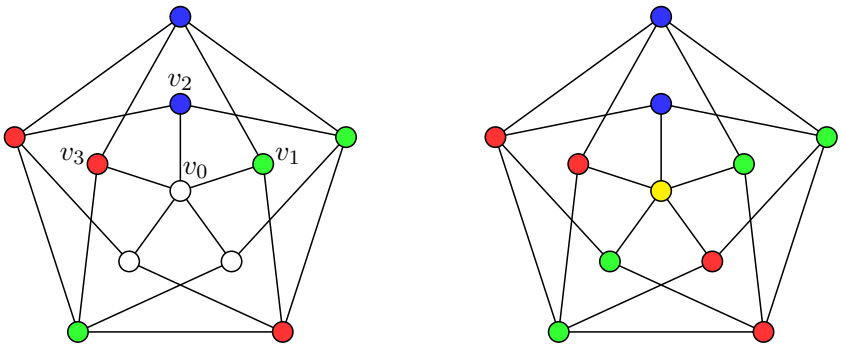
Grötzschev graf G je malo bolj težaven. Velja namreč $\omega(G) = 2$ in $\Delta(G) = 5$. Po trditvi 9.21 in izreku 9.23 lahko sklepamo, da je $2 \leq \chi(G) \leq 5$. Ker G ravno tako ni dvodelen graf (saj očitno vsebuje petcikel), je $\chi(G) \geq 3$.

Velja torej $3 \leq \chi(G) \leq 5$, še vedno imamo tri možnosti za kromatično število.

Poskusimo poiskati 3-barvanje c_3 grafa G. Zaradi petkotniške simetrije grafa G smemo privzeti, da se barve "zunanjih" točk ujemajo z odlikovanim barvanjem s slike 9.32. Posledično mora barvanje c_3 za točke v_1, v_2 in v_3 po vrsti uporabiti zeleno, modro in rdečo barvo, glej sliko 9.34(levo). Toda nobena od omenjenih treh barv ni dopustna v točki v_0, ki je soseda z v_1, v_2 in z v_3. Grötzschev graf torej ni 3-obarvljiv.

Neuspešen poskus iskanja 3-barvanja enostavno razširimo do 4-barvanja grafa G, glej sliko 9.34(desno). Pokazali smo torej, da Grötzschev graf ne dopušča 3-barvanja, ima pa 4-barvanje točk. Torej je $\chi(G) = 4$.

Na tem mestu naj omenimo, da ima G veliko različnih 4-barvanj. Nikakor ne moremo sklepati, da je (edina) točka stopnje 5 v vsakem 4-barvanju grafa G obarvana z "eksklu-



Slika 9.34: Grötzschev graf G ima kromatično število 4.

zivno" barvo.

S tem zgledom zaključujemo razdelek o barvanjih grafov, kot tudi celo grafovsko poglavje.